OE LIBRARY

18

DEC 3 – 1935

Vol. XVIII, No. 6 *PSYCHOLOGICAL REVIEW PUBLICATIONS* December, 1935

Journal of Experimental Psychology

EDITED BY
SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

WALTER S. HUNTER, CLARK UNIVERSITY (*Psychol. Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Psychol. Rev.*)
JOHN A. McGEOCH, WESLEYAN UNIVERSITY (*Psychol. Bull.*)
JOHN F. DASHIELL, UNIV. OF NORTH CAROLINA (*Psychol. Monog.*)

CONTENTS

Studies of Interference in Serial Verbal Reactions: J. R. STROOP 643

An Experimental Test of the Theory of Associative Interference: J. W. NAGGE 663

A Study of Latent Learning: M. T. EATON 683

Certain Somatic Activities in Relation to Successful and Unsuccessful Problem Solving: M. S. CLITES 708

The Effect of a Change of Background on Recall and Relearning: S. G. DULSKY 725

Time-Order Errors in the Discrimination of Short Tonal Durations: L. H. STOTT 741

Interpolation Effects with Different Time Intervals: J. G. NEEDHAM 767

Bilateral Transfer of the Conditioned Knee-Jerk: J. J. GIBSON AND L. HUDSON 774

Macrophonic Speech: E. W. SCRIPTURE 784

The Comparative Value of Continuous Versus Interrupted Sleep: R. W. HUSBAND 792

Distribution of Effort in a Perceptual-Motor Task: E. B. SKAGGS 797

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Review*)
JOHN A. McGEOCH, WESLEYAN UNIVERSITY (*Bulletin*)
JOHN F. DASHIELL, UNIV. OF NORTH CAROLINA (*Monographs*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, June, September, December, March, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). Index: $4.00 per volume.
Journal: $7.00 (Foreign, $7.25). Monographs: $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). Abstracts: $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: Journal, $1.25; Review, $1.00; Abstracts, 75c; Bulletin, 60c.

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp.: $11.00 (Foreign, $11.50).
Bulletin and J. Exp.: $12.00 (Foreign, $12.50).
Review, Bulletin, and J. Exp.: $16.00 (Foreign, $16.75).
Review, Bulletin, J. Exp., and Index: $19.00 (Foreign, $19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XVIII, No. 6 DECEMBER, 1935

STUDIES OF INTERFERENCE IN SERIAL VERBAL REACTIONS

BY J. RIDLEY STROOP*

George Peabody College

INTRODUCTION

Interference or inhibition (the terms seem to have been used almost indiscriminately) has been given a large place in experimental literature. The investigation was begun by the physiologists prior to 1890 (Bowditch and Warren, J. W., 1890) and has been continued to the present, principally by psychologists (Lester, 1932). Of the numerous studies that have been published during this period only a limited number of the most relevant reports demand our attention here.

Münsterberg (1892) studied the inhibiting effects of changes in common daily habits such as opening the door of his room, dipping his pen in ink, and taking his watch out of his pocket. He concluded that a given association can function automatically even though some effect of a previous contrary association remains.

Müller and Schumann (1894) discovered that more time

* The writer wishes to acknowledge the kind assistance received in the preparation of this thesis. He is indebted to Dr. Joseph Peterson for encouragement, helpful suggestions, and criticism of the manuscript; to Major H. W. Fenker, a graduate student in psychology, for helpful suggestions relative to preparation of the manuscript; to Drs. J. Peterson, S. C. Garrison, M. R. Schneck, J. E. Caster, O. A. Simley, W. F. Smith, and to Miss M. Nichol for aid in securing subjects; to some three hundred college students who served as subjects; and to William Fitzgerald of The Peabody Press for substantial assistance in the printing of the test materials.

was necessary to relearn a series of nonsense syllables if the stimulus syllables had been associated with other syllables in the meantime. From their results they deduced the law of associative inhibition which is quoted by Kline (1921, p. 270) as follows: "If *a* is already connected with *b*, then it is difficult to connect it with *k*, *b* gets in the way." Nonsense syllables were also used by Shepard and Fogelsonger (1913) in a series of experiments in association and inhibition. Only three subjects were used in any experiment and the changes introduced to produce the inhibition were so great in many cases as to present novel situations. This latter fact was shown by the introspections. The results showed an increase in time for the response which corresponded roughly to the increase in the complexity of the situation. The only conclusion was stated thus: "We have found then that in acquiring associations there is involved an inhibitory process which is not a mere result of divided paths but has some deeper basis yet unknown" (p. 311).

Kline (1921) used 'meaningful' material (states and capitals, counties and county seats, and books and authors) in a study of interference effects of associations. He found that if the first associative bond had a recall power of 10 percent or less it facilitated the second association, if it had a recall power of 15 percent to 40 percent the inhibitory power was small, if it had a recall power of 45 percent to 70 percent the inhibiting strength approached a maximum, if the recall power was 70 percent to 100 percent the inhibition was of medium strength and in some cases might disappear or even facilitate the learning of a new association.

In card sorting Bergström (1893 and 1894), Brown (1914), Bair (1902), and Culler (1912) found that changing the arrangement of compartments into which cards were being sorted produced interference effects. Bergström (1894, p. 441) concluded that "the interference effect of an association bears a constant relation to the practice effect, and is, in fact, equivalent to it." Both Bair and Culler found that the interference of the opposing habits disappeared if the habits were practiced alternately.

Culler (1912), in the paper already referred to, reported two other experiments. In one experiment the subjects associated each of a series of numbers with striking a particular key on the typewriter with a particular finger; then the keys were changed so that four of the numbers had to be written with fingers other than those formerly used to write them. In the other experiment the subjects were trained to react with the right hand to 'red' and with the left hand to 'blue.' Then the stimuli were interchanged. In the former experiment an interference was found which decreased rapidly with practice. In the latter experiment the interference was overbalanced by the practice effect.

Hunter and Yarbrough (1917), Pearce (1917), and Hunter (1922) in three closely related studies of habit interference in the white rat in a T-shaped discrimination box found that a previous habit interfered with the formation of an 'opposite' habit.

Several studies have been published which were not primarily studies of interference, but which employed materials that were similar in nature to those employed in this research, and which are concerned with why it takes more time to name colors than to read color names. Several of these studies have been reviewed recently by Telford (1930) and by Ligon (1932). Only the vital point of these studies will be mentioned here.

The difference in time for naming colors and reading color names has been variously explained. Cattell (1886) and Lund (1927) have attributed the difference to 'practice.' Woodworth and Wells (1911, p. 52) have suggested that, "The real mechanism here may very well be the mutual interference of the five names, all of which, from immediately preceding use, are 'on the tip of the tongue,' all are equally ready and likely to get in one another's way." Brown (1915, p. 51) concluded "that the difference in speed between color naming and word reading does not depend upon practice" but that (p. 34) "the association process in naming simple objects like colors is radically different from the association process in reading printed words."

Garrett and Lemmon (1924, p. 438) have accounted for their findings in these words, "Hence it seems reasonable to say that interferences which arise in naming colors are due not so much to an equal readiness of the color names as to an equal readiness of the color recognitive processes. Another factor present in interference is very probably the present strength of the associations between colors and their names, already determined by past use." Peterson (1918 and 1925) has attributed the difference to the fact that, "One particular response habit has become associated with each word while in the case of colors themselves a variety of response tendencies have developed." (1925, p. 281.) As pointed out by Telford (1930), the results published by Peterson (1925, p. 281) and also those published by Lund (1927, p. 425) confirm Peterson's interpretation.

Ligon (1932) has published results of a 'genetic study' of naming colors and reading color names in which he used 638 subjects from school grades 1 to 9 inclusive. In the light of his results he found all former explanations untenable (He included no examination of or reference to Peterson's data and interpretation.) and proceeded to set up a new hypothesis based upon a three factor theory, a common factor which he never definitely describes and special factors of word reading and color naming. He points out that the common factor is learned but the special factors are organic. He promises further evidence from studies now in progress.

The present problem grew out of experimental work in color naming and word reading conducted in Jesup Psychological Laboratory at George Peabody College For Teachers. The time for reading names of colors had been compared with the time for naming colors themselves. This suggested a comparison of the interfering effect of color stimuli upon reading names of colors (the two types of stimuli being presented simultaneously) with the interfering effect of word stimuli upon naming colors themselves. In other words, if the word 'red' is printed in blue ink how will the interference of the ink-color 'blue' upon reading the printed word 'red' compare with the interference of the

printed word 'red' upon calling the name of the ink-color 'blue?' The increase in time for reacting to words caused by the presence of conflicting color stimuli is taken as the measure of the interference of color stimuli upon reading words. The increase in the time for reacting to colors caused by the presence of conflicting word stimuli is taken as the measure of the interference of word stimuli upon naming colors. A second problem grew out of the results of the first. The problem was, What effect would practice in reacting to the color stimuli in the presence of conflicting word stimuli have upon the reaction times in the two situations described in the first problem?

Experimental

The materials employed in these experiments are quite different from any that have been used to study interference.[1] In former studies the subjects were given practice in responding to a set of stimuli until associative bonds were formed between the stimuli and the desired responses, then a change was made in the experimental 'set up' which demanded a different set of responses to the same set of stimuli. In the present study pairs of conflicting stimuli, both being inherent aspects of the same symbols, are presented simultaneously (a name of one color printed in the ink of another color—a word stimulus and a color stimulus). These stimuli are varied in such a manner as to maintain the potency of their interference effect. Detailed descriptions of the materials used in each of the three experiments are included in the reports of the respective experiments.

Experiment I

The Effect of Interfering Color Stimuli Upon Reading Names of Colors Serially

Materials:

When this experiment was contemplated, the first task was to arrange suitable tests. The colors used on the Woodworth-Wells color-sheet were considered but two changes were deemed advisable. As the word test to be used in comparison with the

[1] Descoeudres (1914) and also Goodenough and Brian (1929) presented color and form simultaneously in studying their relative values as stimuli.

color test was to be printed in black it seemed well to substitute another color for black as an interfering stimulus. Also, because of the difficulty of printing words in yellow that would approximate the stimulus intensity of the other colors used, yellow was discarded. After consulting with Dr. Peterson, black and yellow were replaced by brown and purple. Hence, the colors used were red, blue, green, brown, and purple. The colors were arranged so as to avoid any regularity of occurrence and so that each color would appear twice in each column and in each row, and that no color would immediately succeed itself in either column or row. The words were also arranged so that the name of each color would appear twice in each line. No word was printed in the color it named but an equal number of times in each of the other four colors; *i.e.* the word 'red' was printed in blue, green, brown, and purple inks; the word 'blue' was printed in red, green, brown, and purple inks; etc. No word immediately succeeded itself in either column or row. The test was printed from fourteen point Franklin lower case type. The word arrangement was duplicated in black print from same type. Each test was also printed in the reverse order which provided a second form. The tests will be known as "Reading color names where the color of the print and the word are different" (*RCNd*),[2] and "Reading color names printed in black" (*RCNb*).

Subjects and Procedure:

Seventy college undergraduates (14 males and 56 females) were used as subjects. Every subject read two whole sheets (the two forms) of each test at one sitting. One half of the subjects of each sex, selected at random, read the tests in the order *RCNb* (form 1), *RCNd* (form 2), *RCNd* (form 1) and *RCNb* (form 2), while the other half reversed the order thus equating for practice and fatigue on each test and form. All subjects were seated so as to have good daylight illumination from the left side only. All subjects were in the experimental room a few minutes before beginning work to allow the eyes to adjust to light conditions. The subjects were volunteers and apparently the motivation was good.

A ten-word sample was read before the first reading of each test. The instructions were to read as quickly as possible and to leave no errors uncorrected. When an error was left the subject's attention was called to that fact as soon as the sheet was finished. On the signal "Ready! Go!" the sheet which the subject held face down was turned by the subject and read aloud. The words were followed on another sheet (in black print) by the experimenter and the time was taken with a stop watch to a fifth of a second. Contrary to instructions 14 subjects left a total of 24 errors uncorrected on the *RCNd* test, 4 was the maximum for any subject, and 4 other subjects left 1 error each on the *RCNb* test. As each subject made 200 reactions on each test this small number of errors was considered negligible. The work was done under good daylight illumination.

Results: Table I gives the means (m), standard deviations (σ), differences (D), probable error of the difference (PE_d), and the reliability of the difference (D/PE_d) for the whole group and for each sex.

Observation of the bottom line of the table shows that it

[2] In Appendix *A* will be found a key to all symbols and abbreviations used in this paper.

TABLE I

THE MEAN TIME IN SECONDS FOR READING ONE HUNDRED NAMES OF COLORS PRINTED IN COLORS DIFFERENT FROM THAT NAMED BY THE WORD AND FOR ONE HUNDRED NAMES OF COLORS PRINTED IN BLACK

Sex	No. Ss.	*RCNd*	σ	*RCNb*	σ	*D*	PE_d	D/PE_d
Male	14	43.20	4.98	40.81	4.97	2.41	1.27	1.89
Female	56	43.32	6.42	41.04	4.78	2.28	.72	3.16
Male and Female	70	43.30	6.15	41.00	4.84	2.30	.63	3.64

took an average of 2.3 seconds longer to read 100 color names printed in colors different from that named by the word than to read the same names printed in black. This difference is not reliable which is in agreement with Peterson's prediction made when the test was first proposed. The means for the sex groups show no particular difference. An examination of the means and standard deviations for the two tests shows that the interference factor caused a slight increase in the variability for the whole group and for the female group, but a slight decrease for the male group.

Table II presents the same data arranged on the basis

TABLE II

SHOWING DATA OF TABLE I ARRANGED ON THE BASIS OF COLLEGE CLASSIFICATION

College Year	No. Ss.	*RCNd*	σ	*RCNb*	σ	*D*	D/PE_d
1st	35	43.9	6.31	41.7	5.58	2.2	.38
2d	20	44.9	6.74	41.8	4.32	3.1	.57
3d	8	39.8	4.62	39.2	3.73	.6	.16
4th	7	40.8	3.60	39.2	2.93	1.6	.51

of college classification. Only college years one and two contain a sufficient number of cases for comparative purposes. They show no differences that approach reliability.

EXPERIMENT 2

The Effect of Interfering Word Stimuli upon Naming Colors Serially

Materials:

For this experiment the colors of the words in the *RCNd* test, described in Experiment 1, were printed in the same order but in the form of solid squares (■) from 24 point type instead of words. This sort of problem will be referred to as the

color test was to be printed in black it seemed well to substitute another color for black as an interfering stimulus. Also, because of the difficulty of printing words in yellow that would approximate the stimulus intensity of the other colors used, yellow was discarded. After consulting with Dr. Peterson, black and yellow were replaced by brown and purple. Hence, the colors used were red, blue, green, brown, and purple. The colors were arranged so as to avoid any regularity of occurrence and so that each color would appear twice in each column and in each row, and that no color would immediately succeed itself in either column or row. The words were also arranged so that the name of each color would appear twice in each line. No word was printed in the color it named but an equal number of times in each of the other four colors; *i.e.* the word 'red' was printed in blue, green, brown, and purple inks; the word 'blue' was printed in red, green, brown, and purple inks; etc. No word immediately succeeded itself in either column or row. The test was printed from fourteen point Franklin lower case type. The word arrangement was duplicated in black print from same type. Each test was also printed in the reverse order which provided a second form. The tests will be known as "Reading color names where the color of the print and the word are different" (*RCNd*),[2] and "Reading color names printed in black" (*RCNb*).

Subjects and Procedure:

Seventy college undergraduates (14 males and 56 females) were used as subjects. Every subject read two whole sheets (the two forms) of each test at one sitting. One half of the subjects of each sex, selected at random, read the tests in the order *RCNb* (form 1), *RCNd* (form 2), *RCNd* (form 1) and *RCNb* (form 2), while the other half reversed the order thus equating for practice and fatigue on each test and form. All subjects were seated so as to have good daylight illumination from the left side only. All subjects were in the experimental room a few minutes before beginning work to allow the eyes to adjust to light conditions. The subjects were volunteers and apparently the motivation was good.

A ten-word sample was read before the first reading of each test. The instructions were to read as quickly as possible and to leave no errors uncorrected. When an error was left the subject's attention was called to that fact as soon as the sheet was finished. On the signal "Ready! Go!" the sheet which the subject held face down was turned by the subject and read aloud. The words were followed on another sheet (in black print) by the experimenter and the time was taken with a stop watch to a fifth of a second. Contrary to instructions 14 subjects left a total of 24 errors uncorrected on the *RCNd* test, 4 was the maximum for any subject, and 4 other subjects left 1 error each on the *RCNb* test. As each subject made 200 reactions on each test this small number of errors was considered negligible. The work was done under good daylight illumination.

Results: Table I gives the means (m), standard deviations (σ), differences (D), probable error of the difference (PE_d), and the reliability of the difference (D/PE_d) for the whole group and for each sex.

Observation of the bottom line of the table shows that it

[2] In Appendix *A* will be found a key to all symbols and abbreviations used in this paper.

TABLE I

THE MEAN TIME IN SECONDS FOR READING ONE HUNDRED NAMES OF COLORS PRINTED IN COLORS DIFFERENT FROM THAT NAMED BY THE WORD AND FOR ONE HUNDRED NAMES OF COLORS PRINTED IN BLACK

Sex	No. Ss.	*RCNd*	σ	*RCNb*	σ	*D*	PE_d	D/PE_d
Male	14	43.20	4.98	40.81	4.97	2.41	1.27	1.89
Female	56	43.32	6.42	41.04	4.78	2.28	.72	3.16
Male and Female	70	43.30	6.15	41.00	4.84	2.30	.63	3.64

took an average of 2.3 seconds longer to read 100 color names printed in colors different from that named by the word than to read the same names printed in black. This difference is not reliable which is in agreement with Peterson's prediction made when the test was first proposed. The means for the sex groups show no particular difference. An examination of the means and standard deviations for the two tests shows that the interference factor caused a slight increase in the variability for the whole group and for the female group, but a slight decrease for the male group.

Table II presents the same data arranged on the basis

TABLE II

SHOWING DATA OF TABLE I ARRANGED ON THE BASIS OF COLLEGE CLASSIFICATION

College Year	No. Ss.	*RCNd*	σ	*RCNb*	σ	*D*	D/PE_d
1st	35	43.9	6.31	41.7	5.58	2.2	.38
2d	20	44.9	6.74	41.8	4.32	3.1	.57
3d	8	39.8	4.62	39.2	3.73	.6	.16
4th	7	40.8	3.60	39.2	2.93	1.6	.51

of college classification. Only college years one and two contain a sufficient number of cases for comparative purposes. They show no differences that approach reliability.

EXPERIMENT 2

The Effect of Interfering Word Stimuli upon Naming Colors Serially

Materials:

For this experiment the colors of the words in the *RCNd* test, described in Experiment 1, were printed in the same order but in the form of solid squares (■) from 24 point type instead of words. This sort of problem will be referred to as the

"Naming color test" (*NC*). The *RCNd* test was employed also but in a very different manner from that in Experiment 1. In this experiment the colors of the print of the series of names were to be called in succession ignoring the colors named by the words; *e.g.* where the word 'red' was printed in blue it was to be called 'blue,' where it was printed in green it was to be called 'green,' where the word 'brown' was printed in red it was to be called 'red,' etc. Thus color of the print was to be the controlling stimulus and not the name of the color spelled by the word. This is to be known as the "Naming color of word test where the color of the print and the word are different" (*NCWd*). (See Appendix *B*.)

Subjects and Procedure:

One hundred students (88 college undergraduates, 29 males and 59 females, and 12 graduate students, all females) served as subjects. Every subject read two whole sheets (the two forms) of each test at one sitting. Half of the subjects read in the order *NC*, *NCWd*, *NCWd*, *NC*, and the other half in the order *NCWd*, *NC*, *NC*, *NCWd*, thus equating for practice and fatigue on the two tests. All subjects were seated (in their individual tests) near the window so as to have good daylight illumination from the left side. Every subject seemed to make a real effort.

A ten-word sample of each test was read before reading the test the first time. The instructions were to name the colors as they appeared in regular reading line as quickly as possible and to correct all errors. The methods of starting, checking errors, and timing were the same as those used in Experiment 1. The errors were recorded and for each error not corrected, twice the average time per word for the reading of the sheet on which the error was made was added to the time taken by the stop watch. This plan of correction was arbitrary but seemed to be justified by the situation. There were two kinds of failures to be accounted for: first, the failure to see the error: and second, the failure to correct it. Each phase of the situation gave the subject a time advantage which deserved taking note of. Since no accurate objective measure was obtainable and the number of errors was small the arbitrary plan was adopted. Fifty-nine percent of the group left an average of 2.6 errors uncorrected on the *NCWd* test (200 reactions) and 32 percent of the group left an average of 1.2 errors uncorrected on the *NC* test (200 reactions). The correction changed the mean on the *NCWd* test from 108.7 to 110.3 and the mean of the *NC* test from 63.0 to 63.3.

Results: The means of the times for the *NC* and *NCWd* tests for the whole group and for each sex are presented in Table III along with the difference, the probable error of the

TABLE III

THE MEAN TIME FOR NAMING ONE HUNDRED COLORS PRESENTED IN SQUARES AND IN THE PRINT OF WORDS WHICH NAME OTHER COLORS

Sex	No. Ss.	*NCWd*	σ	*NC*	σ	*D/NC*	*D*	PE_d	D/PE_d
Male.....	29	111.1	21.6	69.2	10.8	.61	42.9	3.00	13.83
Female.....	71	107.5	17.3	61.0	10.5	.76	46.5	1.62	28.81
Male and Female.....	100	110.3	18.8	63.3	10.8	.74	47.0	1.50	31.38

difference, the reliability of the difference, and the difference divided by the mean time for the naming color test.

The comparison of the results for the whole group on the *NC* and *NCWd* test given in the bottom line of the table indicates the strength of the inteference of the habit of calling words upon the activity of naming colors. The mean time for 100 responses is increased from 63.3 seconds to 110.3 seconds or an increase of 74 percent. (The medians on the two tests are 61.9 and 110.4 seconds respectively.) The standard deviation is increased in approximately the same ratio from 10.8 to 18.8. The coefficient of variability remains the same to the third decimal place ($\sigma/m = .171$). The difference between means may be better evaluated when expressed in terms of the variability of the group. The difference of 47 seconds is 2.5 standard deviation units in terms of the *NCWd* test or 4.35 standard deviation units on the *NC* test. The former shows that 99 percent of the group on the *NCWd* test was above the mean on the *NC* test (took more time); and the latter shows that the group as scored on the *NC* test was well below the mean on the *NCWd* test. These results are shown graphically in Fig. 1 where histograms and

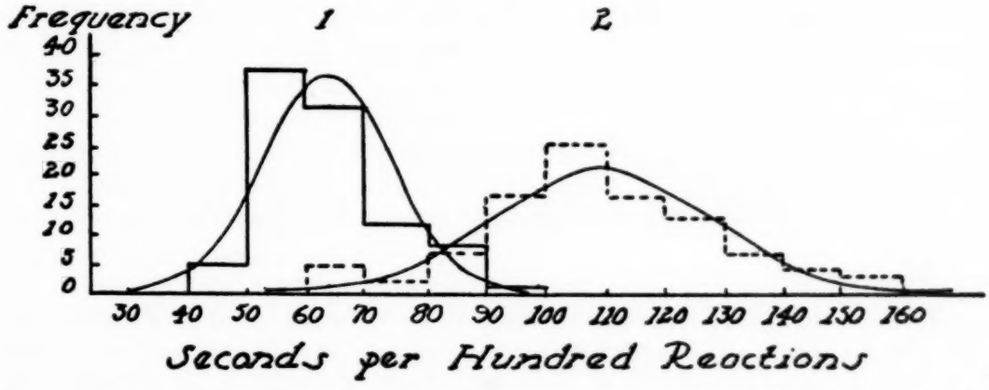

FIG. 1. Showing the effect of interference on naming colors. No interference (1); interference (2).

normal curves (obtained by the Gaussian formula) of the two sets of data are superimposed. The small area in which the curves overlap and the 74 percent increase in the mean time for naming colors caused by the presence of word stimuli show the marked interference effect of the habitual response of calling words.

The means for the sex groups on the *NCWd* test show a difference of 3.6 seconds which is only 1.16 times its probable error; but the means on the *NC* test have a difference of 8.2 seconds which is 5.17 times its probable error. This reliable sex-difference favoring the females in naming colors agrees with the findings of Woodworth-Wells (1911), Brown (1915), Ligon (1932), etc.

The same data are arranged according to college classification in Table IV. There is some indication of improve-

TABLE IV

SHOWING THE DATA OF TABLE III ARRANGED ON THE BASIS OF COLLEGE CLASSIFICATION

Class	No. Ss.	*NCWd*	σ	*NC*	σ	*D*	*D/NC*	D/PE_d
1st yr.	17	116.5	24.9	70.9	15.9	45.6	.64	22.7
2d yr.	37	114.4	18.0	66.1	10.6	48.3	.73	32.6
3d yr.	12	106.1	14.0	62.8	7.0	43.3	.69	41.2
4th yr.	22	96.6	16.8	57.8	8.9	38.8	.67	30.3
Graduates.	12	111.2	19.4	59.9	11.5	51.3	.86	37.6

ment of the speed factor for both tests as the college rank improves. The relative difference between the two tests, however, remains generally the same except for fluctuations which are probably due to the variation in the number of cases.

EXPERIMENT 3

The Effects of Practice upon Interference

Materials:

The tests used were the same in character as those described in Experiments 1 and 2 (*RCNb*, *RCNd* ,*NC*, and *NCWd*) with some revision. The *NC* test was printed in swastikas (卐) instead of squares (■). Such a modification allowed white to appear in the figure with the color, as is the case when the color is presented in the printed word. This change also made it possible to print the *NC* test in shades which more nearly match those in the *NCWd* test. The order of colors was determined under one restriction other than those given in section 2. Each line contained one color whose two appearances were separated by only one other color. This was done to equate, as much as possible, the difficulty of the different lines of the test so that any section of five lines would approximate the difficulty of any other section of five lines. Two forms of the tests were printed; in one the order was the inverse of that in the other.

Subjects and Procedure:

Thirty-two undergraduates in the University of Arizona (17 males and 15 females), who offered their services, were the subjects. At each day's sitting 4 half-sheets of the same test were read, and the average time (after correction was made for errors according to the plan outlined in Experiment 2) was recorded as the day's score. Only a few errors were left uncorrected. The largest correction made on the practice test changed the mean from 49.3 to 49.6. The plan of experimentation was as follows:

Day	1	2	3	4	5	6	7
Test	*RCNb*	*RCNd*	*NC*	*NCWd*	*NCWd*	*NCWd*	*NCWd*
Day	8	9	10	11	12	13	14
Test	*NCWd*	*NCWd*	*NCWd*	*NCWd*	*NC*	*RCNd*	*RCNd*

On the 1st day the *RCNb* test was used to acquaint the subjects with the experimental procedure and improve the reliability of the 2d day's test. The *RCNd* test was given the 2d day and the 13th day to obtain a measure of the interference developed by practice on the *NC* and *NCWd* tests. The *RCNd* test was given the 14th day to get a measure of the effect of a day's practice upon the newly developed interference. The *NC* test was given the 3d and 12th days, just before and just after the real practice series, so that actual change in interference on the *NCWd* test might be known. The test schedule was followed in regular daily order with two exceptions. There were two days between test days 3 and 4, and also two between test days 8 and 9, in which no work was done. These irregularities were occasioned by week-ends. Each subject was assigned a regular time of day for his work throughout the experiment. All but two subjects followed the schedule with very little irregularity. These two were finally dropped from the group and their data rejected.

All of the tests were given individually by the author. The subject was seated near a window so as to have good daylight illumination from the left side. There was no other source of light. Every subject was in the experimental room a few minutes before beginning work to allow his eyes to adapt to the light conditions. To aid eye-adaptation and also to check for clearness of vision each subject read several lines in a current magazine. Every subject was given Dr. Ishihara's test for color vision. One subject was found to have some trouble with red-green color vision; and her results were discarded though they differed from others of her sex only in the number of errors made and corrected.

Results: The general results for the whole series of tests are shown in Table V which presents the means, standard deviations, and coefficients of variability for the whole group and for each sex separately, together with a measure of sex differences in terms of the probable error of the difference. Table VI, which is derived from Table V, summarizes the practice effects upon the respective tests. The graphical representation of the results in the practice series gives the learning curve presented in Fig. 2.

TABLE V

SHOWING THE EFFECTS OF PRACTICE ON THE *NCWd* TEST UPON ITSELF, UPON THE *NC* TEST, AND UPON THE *RCNd* TEST IN TERMS OF MEAN SCORES, STANDARD DEVIATIONS (σ), AND COEFFICIENTS OF VARIABILITY (σ/m) FOR THIRTY-TWO COLLEGE STUDENTS

The score is the average time for four trials of fifty reactions each.

Sex	No. Ss.	Initial Tests						Days of Practice on the *NCWd* Test							
		RCNb	σ	*RCNd*	σ	*NC*	σ	1	σ	2	σ	3	σ	4	σ
Male	17	19.8	1.8	19.6	2.5	30.6	3.6	51.2	8.5	41.6	7.8	38.2	7.6	37.3	8.0
Female	15	18.3	2.9	19.1	3.4	26.5	2.8	47.8	4.2	39.1	4.4	35.8	3.4	33.7	3.7
M & F	32	19.1	2.6	19.4	3.0	28.7	3.5	49.6	7.1	40.5	6.4	37.1	6.1	35.7	6.5
		Sex Differences													
M & F		1.5		.5		4.1		3.4		2.5		2.4		3.6	
PE_d		.49		.70		.76		1.55		1.47		1.36		1.45	
D/PE_d		3.06		.71		5.39		2.19		1.70		1.76		2.48	
Sex		Coefficients of Variability													
Male		.09±.011		.13±.015		.12±.014		.17±.020		.19±.022		.20±.024		.22±.026	
Female		.16±.024		.18±.028		.11±.016		.09±.013		.11±.017		.09±.014		.11±.017	
M & F		.14±.012		.15±.013		.12±.010		.14±.012		.16±.014		.17±.014		.18±.016	

Sex	No. Ss.	Days of Practice on the *NCWd* Test								Final Tests					
		5	σ	6	σ	7	σ	8	σ	*NC*	σ	*RCNd*	σ	*RCNd*	σ
Male	17	36.3	7.4	33.9	7.3	33.5	6.7	33.4	7.1	25.9	4.2	37.3	13.7	22.2	4.8
Female	15	32.8	4.3	32.3	4.0	31.6	3.3	31.5	3.3	23.6	1.9	32.0	6.2	21.8	6.1
M & F	32	34.9	6.2	33.2	5.4	32.6	5.5	32.8	6.1	24.7	3.2	34.8	11.7	22.0	5.5
		Sex Differences													
M & F		3.5		1.6		1.9		1.9		2.3		5.3		.4	
PE_d		1.41		1.34		1.23		1.30		.77		2.56		1.31	
D/PE_d		2.48		1.19		1.54		1.46		2.99		2.07		.31	
Sex		Coefficients of Variability													
Male		.20±.024		.22±.026		.20±.024		.21±.025		.16±.019		.37±.048		.22±.026	
Female		.13±.020		.12±.019		.10±.016		.11±.016		.08±.012		.19±.030		.28±.045	
M & F		.18±.016		.16±.014		.17±.015		.19±.015		.13±.011		.34±.031		.25±.022	

The Effect of Practice on the NCWd Test upon Itself

The data to be considered here are those given in the section of Table V under the caption "Days of Practice on the *NCWd* Test." They are also presented in summary in the left section of Table VI and graphically in Fig. 2. From all

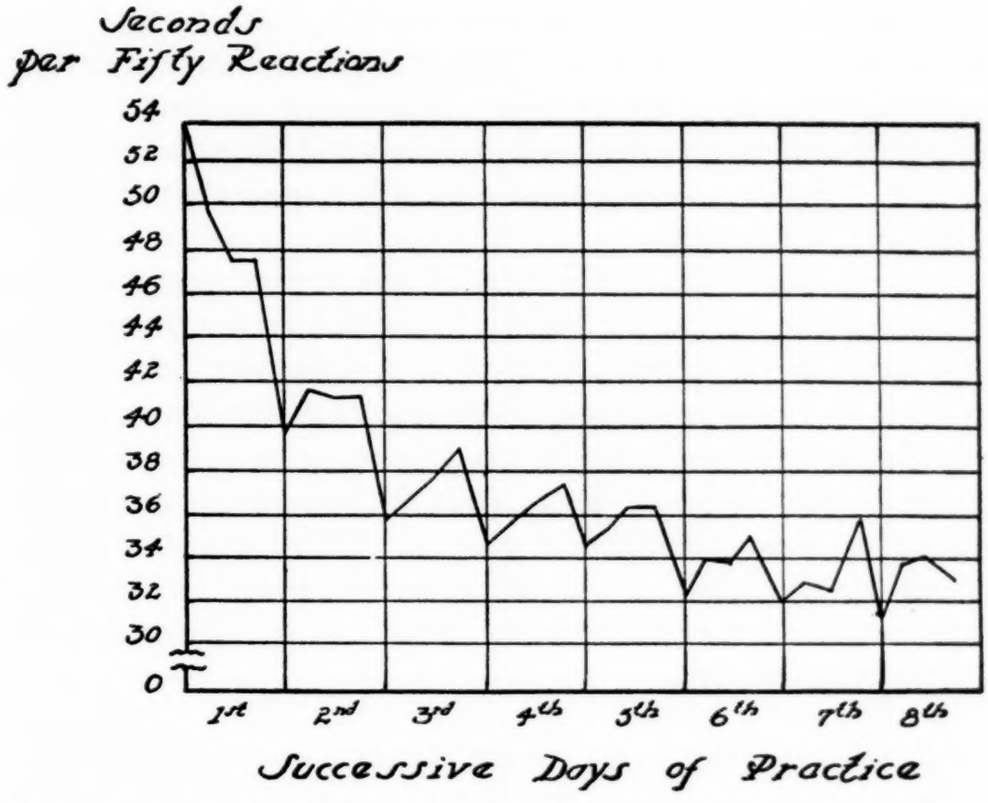

FIG. 2. Mean scores for the group in each of the four half-sheets of the *NCWd* test, which constituted the daily practice.

three presentations it is evident that the time score is lowered considerably by practice. Reference to Table VI shows a gain of 16.8 seconds or 33.9 percent of the mean of the 1st day's practice. The practice curve is found to resemble very much the 'typical' learning curve when constructed on

TABLE VI

A SUMMARY OF THE MEANS IN TABLE V, SHOWING THE EFFECT OF PRACTICE IN THE *NCWd* TEST UPON THE *NCWd*, THE *NC*, AND THE *RCNd* TESTS

Test	*NCWd*			*NC*			*RCNd*		
Sex	M	F	M & F	M	F	M & F	M	F	M & F
Initial Score	51.2	47.8	49.6	30.6	26.5	28.7	19.6	19.1	19.4
Final Score	33.4	31.5	32.8	25.9	23.6	24.7	37.3	32.0	34.8
Gain	17.8	16.3	16.8	4.7	2.9	4.0	−17.7	−12.9	−15.4
Percent Gain	34.8	34.1	33.9	15.4	10.9	13.9	−90.3	−67.5	−79.3

Minus sign shows loss

time units. The coefficient of variability is increased from .14 ± .012 to .19 ± .015. This difference divided by its probable error gives 2.60 which indicates that it is not reliable. The probability of a real increase in variability, however, is 24 to 1. Hence, practice on the *NCWd* test serves to increase individual differences.

An examination of the data of the sex groups reveals a difference in speed on the *NCWd* test which favors the females. This is to be expected as there is a difference in favor of females in naming colors. Though the difference is not reliable in any one case it exists throughout the practice series; indicating that the relative improvement is approximately the same for the two groups. This latter fact is also shown by the ratio of the difference between the halves of practice series to the first half. It is .185 for the males and .180 for the females.

The Effect of Practice on the NCWd Test upon the NC Test

The middle section of Table VI shows a gain on the *NC* test of 4.0 seconds or 13.9 percent of the initial score. This is only 23.7 percent of the gain on the *NCWd* test which means that less than one fourth of the total gain on the *NCWd* test is due to increase in speed in naming colors. The improvement is greater for the males, which is accounted for by the fact that there is more difference between naming colors and reading names of colors for the males than for the females.

The Effects in the RCNd Test of Practice on the NCWd and NC Tests

The right section of Table VI shows that the practice on the *NCWd* and *NC* tests resulted in heavy loss in speed on the *RCNd* test. A comparison of the right and left sections of the table shows that the loss on the *RCNd* test, when measured in absolute units, is practically equal to the gain on the *NCWd* test; when measured in relative units it is much greater. It is interesting to find that in ten short practice periods the relative values of opposing stimuli can be modified so greatly.

There is little relation, however, between the gain in one case and the loss in the other. The correlation between gain and loss in absolute units is .262 ± .11, while the correlation between percent of gain and percent of loss is .016 ± .17, or zero. This is what one might expect.

From a consideration of the results of the two applications of the *RCNd* test given in the final tests of Table V it is evident that the newly developed interference disappears very rapidly with practice. From one day to the next the mean decreases from 34.8 to 22.0 seconds. This indicates that renewing the effectiveness of old associations which are being opposed by newly formed ones is easier than strengthening new associations in opposition to old well established ones.

The variability of the group is increased by the increase in interference due to practice on the *NCWd* test. The coefficient of variability increases from .15 ± .013 to .34 ± .031, the difference divided by its probable error being 5.65. This is not surprising as the degree of the interference varies widely for different subjects. Its degree is determined by the learning on the practice series which is shown by the individual results to vary considerably. One day's practice on the *RCNd* test reduced the variability from .34 ± .031 to .25 ± .022. The decrease in variability is 2.3 times its probable error.

The data from this experiment present interesting findings on the effect of practice upon individual differences. The results which have already been discussed separately are presented for comparison in Table VII.

TABLE VII

THE EFFECTS OF PRACTICE ON THE *NCWd* TEST AND THE *RCNd* TEST UPON THE COEFFICIENT OF VARIABILITY FOR THE GROUP

Test	No. Ss.	Coefficients of Variability		D	PE_d	D/PE_d
		Initial	Final			
NCW	32	.14	.19	.05	.034	2.60
RCNd	32	.34	.25	.09	.037	2.33

These results show that practice increases individual differences where a stimulus to which the subjects have an habitual reaction pattern is interfering with reactions to a stimulus for which the subjects do not have an habitual reaction pattern (the word stimulus interfering with naming colors, *NCWd* test); but decreases individual differences where a stimulus to which the subjects do not have an habitual reaction pattern is interfering with reactions to a stimulus for which the subjects have an habitual reaction pattern (the color stimulus interfering with reading words—*RCNd* test). There are two other variables involved, however: initial variability and length of practice. Thus in the *NCWd* test the initial variability was less, the difficulty greater, and the practice greater than in the *RCNd* test. These findings lend some support to Peterson's hypothesis, "Subjects of normal heterogeneity would become more alike with practice on the simpler processes or activities, but more different on the more complex activities" (Peterson and Barlow, 1928, p. 228).

A sex difference in naming colors has been found by all who have studied color naming and has been generally attributed to the greater facility of women in verbal reactions than of men. There is some indication in our data that this sex difference may be due to the difference in the accustomed reaction of the two sexes to colors as stimuli. In other words responding to a color stimulus by naming the color may be more common with females than with males. This difference is probably built up through education. Education in color is much more intense for girls than for boys as observing, naming, and discussing colors relative to dress is much more common among girls than among boys. The practice in naming colors in the *NCWd* test decreased the difference between the sex groups on the *NC* test from a difference 5.38 times its probable error to a difference 2.99 times its probable error. This decrease in the difference due to practice favors the view that the difference has been acquired and is therefore a product of training.

SUMMARY

1. Interference in serial verbal reactions has been studied by means of newly devised experimental materials. The source of the interference is in the materials themselves. The words red, blue, green, brown, and purple are used on the test sheet. No word is printed in the color it names but an equal number of times in each of the other four colors; *i.e.* the word 'red' is printed in blue, green, brown, and purple inks; the word 'blue' is printed in red, green, brown, and purple inks; etc. Thus each word presents the name of one color printed in ink of another color. Hence, a word stimulus and a color stimulus both are presented simultaneously. The words of the test are duplicated in black print and the colors of the test are duplicated in squares or swastikas. The difference in the time for reading the words printed in colors and the same words printed in black is the measure of the interference of color stimuli upon reading words. The difference in the time for naming the colors in which the words are printed and the same colors printed in squares (or swastikas) is the measure of the interference of conflicting word stimuli upon naming colors.

2. The interference of conflicting color stimuli upon the time for reading 100 words (each word naming a color unlike the ink-color of its print) caused an increase of only 2.3 seconds or 5.6 percent over the normal time for reading the same words printed in black. This increase is not reliable. But the interference of conflicting word stimuli upon the time for naming 100 colors (each color being the print of a word which names another color) caused an increase of 47.0 seconds or 74.3 percent of the normal time for naming colors printed in squares.

These tests provide a unique basis (the interference value) for comparing the effectiveness of the two types of associations. Since the presence of the color stimuli caused no reliable increase over the normal time for reading words ($D/PE_d = 3.64$) and the presence of word stimuli caused a considerable increase over the normal time for naming colors (4.35 standard deviation units) the associations that have been

formed between the word stimuli and the reading response are evidently more effective than those that have been formed between the color stimuli and the naming response. Since these associations are products of training, and since the difference in their strength corresponds roughly to the difference in training in reading words and naming colors, it seems reasonable to conclude that the difference in speed in reading names of colors and in naming colors may be satisfactorily accounted for by the difference in training in the two activities. The word stimulus has been associated with the specific response 'to read,' while the color stimulus has been associated with various responses: 'to admire,' 'to name,' 'to reach for,' 'to avoid,' etc.

3. As a test of the permanency of the interference of conflicting word stimuli to naming colors eight days practice (200 reactions per day) were given in naming the colors of the print of words (each word naming a color unlike the ink-color of its print). The effects of this practice were as follows: 1. It decreased the interference of conflicting word stimuli to naming colors but did not eliminate it. 2. It produced a practice curve comparable to that obtained in many other learning experiments. 3. It increased the variability of the group. 4. It shortened the reaction time to colors presented in color squares. 5. It increased the interference of conflicting color stimuli upon reading words.

4. Practice was found either to increase or to decrease the variability of the group depending upon the nature of the material used.

5. Some indication was found that the sex difference in naming colors is due to the difference in the training of the two sexes.

(Manuscript received August 15, 1934)

References

Bair, J. H., The practice curve: A study of the formation of habits. *Psychol. Rev. Monog. Suppl.*, 1902 (No. 19), 1–70.

Bergström, J. A., Experiments upon physiological memory. *Amer. J. Psychol.*, 1893, **5**, 356–359.

Bergström, J. A., The relation of the interference of the practice effect of an association. *Amer. J. Psychol.*, 1894, **6**, 433–442.

BOWDITCH, H. P., AND WARREN, J. W., The knee-jerk and its physiological modifications. *J. Physiology*, 1890, **11**, 25–46.

BROWN, WARNER, Practice in associating color names with colors. *Psychol. Rev.*, 1915, **22**, 45–55.

BROWN, WARNER, Habit interference in card sorting. *Univ. of Calif. Studies in Psychol.*, 1914, V. i, No. 4.

CATTELL, J. McK., The time it takes to see and name objects. *Mind*, 1886, **11**, 63–65.

CULLER, A. J., Interference and adaptability. *Arch. of Psychol.*, 1912, **3** (No. 24), 1–80.

DESCOEUDRES, A., Couleur, forme, ou nombre. *Arch. de psychol.*, 1914, **14**, 305–341.

GARRETT, H. E., AND LEMMON, V. W., An analysis of several well-known tests. *J. Appld. Psychol.*, 1924, **8**, 424–438.

GOODENOUGH, F. L., AND BRIAN, C. R., Certain factors underlying the acquisition of motor skill by pre-school children. J. EXPER. PSYCHOL., 1929, **12**, 127–155.

HUNTER, W. S., AND YARBROUGH, J. U., The interference of auditory habits in the white rat. *J. Animal Behav.*, 1917, **7**, 49–65.

HUNTER, W. S., Habit interference in the white rat and in the human subject. *J. Comp. Psychol.*, 1922, **2**, 29–59.

KLINE, L. W., An experimental study of associative inhibition. J. EXPER. PSYCHOL., 1921, **4**, 270–299.

LESTER, O. P., Mental set in relation to retroactive inhibition. J. EXPER. PSYCHOL., 1932, **15**, 681–699.

LIGON, E. M. A., Genetic study of color naming and word reading. *Amer. J. Psychol.*, 1932, **44**, 103–121.

LUND, F. H., The role of practice in speed of association. J. EXPER. PSYCHOL., 1927, **10**, 424–433.

MÜLLER, G. E., AND SCHUMANN, F., Experimentelle Beiträge zu Untersuchung des Gedächtnisses. *Zsch. f. Psychol.*, 1894, **6**, 81–190.

MÜNSTERBERG, HUGO, Gedächtnisstudien. *Beiträge zur Experimentellen Psychologie*, 1892, **4**, 70.

PEARCE, BENNIE D., A note on the interference of visual habits in the white rat. *J. Animal Behav.*, 1917, **7**, 169–177.

PETERSON, J., AND BARLOW, M. C., *The effects of practice on individual differences.* The 27th Year Book of Nat. Soc. Study of Educ., Part II, 1928, 211–230.

PETERSON, J., LANIER, L. H., AND WALKER, H. M., Comparisons of white and negro children. *J. Comp. Psychol.*, 1925, **5**, 271–283.

PETERSON, J., AND DAVID, Q. J., *The psychology of handling men in the army.* Minneapolis, Minn. Perine Book Co., 1918, pp. 146.

SHEPARD, J. F., AND FOGELSONGER, H. M., Association and inhibition. *Psychol. Rev.*, 1913, **20**, 291–311.

TELFORD, C. W., Differences in responses to colors and their names. *J. Genet. Psychol.*, 1930, **37**, 151–159.

WOODWORTH, R. S., AND WELLS, F. L., Association tests. *Psychol. Rev. Monog. Suppl.*, 1911, **13** (No. 57), pp. 85.

Appendix *A*

A Key to Symbols and Abbreviations

NC	Naming Colors.
NCWd	Naming the Colors of the Print of Words Where the Color of the Print and the Word are Different.
RCNb	Reading Color Names Printed in Black Ink.
RCNd	Reading Color Names Where the Color of the Print and the Word are Different.
D	Difference.
D/PE_d	Difference divided by the probable error of the difference.
M & F	Males and Females.
PE_d	Probable error of the difference.
σ	Sigma or standard deviation.
σ/m	Standard deviation divided by the mean.

AN EXPERIMENTAL TEST OF THE THEORY OF ASSOCIATIVE INTERFERENCE

BY JOSEPH W. NAGGE

University of Chicago

Introduction

The experimental work on retroactive inhibition has led to the formulation of two main explanatory hypotheses. The first of these is the Perseveration Theory formulated by Müller and Pilzecker (**17**). The second is that of Associative Interference.

The Perseveration Theory rested upon the assumption that, after learning, the neural elements involved therein continued to be automatically excited for a period of time. This perseveration process was held to be essential to a proper fixation and any interference with it led to forgetting. Hence the temporal position of the interpolated activity, in relation to the first activity, should be an important factor in the production of retroactive inhibition. The greatest retroactive inhibition should occur when the interpolated material is learned immediately after the original material. Robinson (**21**), Whitely (**26**), and McGeoch (**10**) (**11**) in two studies, have failed to verify the assumption. Skaggs (**23**) offers data which bear out the assumption.

The Associative Interference Theory holds that when interpolated learning occurs, associations are set up between elements of the interpolated and original materials which tend to disrupt the original learning by cross association. The reinstatement of the first activity, after the acquisition of the second, requires the dissociation of elements from the new context and the ultimate recombining of them into their original system of relationships. The Theory of Associative Interference leads to the following deductions: (1) the amount of retroaction should increase as the degree of similarity of the original and the interpolated activities increases; (2)

the amount of retroaction should decrease with the greater degree of integration of the original material; and (3) retroactive inhibition should decrease as the conditions of association between the first and the interpolated activities are lessened through dissociation. The first of these deductions is made upon the assumption that the closer the similarity of the first and the second activities, the greater will be the number of associations common to both and hence the greater the interference. The second deduction follows the assumption that the greater the integration of the material the more firmly will its associative bonds be made, and hence the more resistant it will become to associative interference from the second. The third deduction is based upon the assumption that as the conditions of association between the two activities are lessened, there will be less opportunity for cross associations, and hence, less interference. The first deduction is supported by the work on similarity by Robinson (**21**), Whitely (**26**), Skaggs (**23**), Harden (**6**), Cheng (**3**), McGeoch and McDonald (**13**), and Johnson (**8**). The second deduction is supported by the work of McGeoch (**12**), Robinson and Heron (**22**), and by that of Tolman (**24**). The third deduction has not been adequately tested. Mitchell's (**16**) experiment is the only test of the assumption. Her experiment, however, is inadequate as she presents data from but two subjects and hence makes no statistical treatment. The method of dissociation employed was that of hypnosis. The original and the interpolated learning occurred in the hypnotic or in the waking states. Altogether there were four conditions of the experiment, two dissociated and two control conditions. Those .conditions in which the original and the interpolated learning occurred in different states showed no advantage in retention over the control conditions in which both activities proceeded in similar states. Rather the dissociated state conditions showed a lower retention than the control conditions. Further and more adequate tests should be made of this third assumption. It is the purpose of the following experiments to further test this assumption.

Experiments

Problem

It is the purpose of the experiments which follow to further test the third assumption of the Theory of Associative Interference, that retroactive inhibition should decrease as the conditions of association between the first and the interpolated activities are lessened through dissociation. The Theory of Associative Interference holds that, in the fixation of the interpolated activity, some of the associative connections established in the original activity are somehow cut across and new associations formed. It is these cross associations which interfere with the retention of the first activity. It follows from the theory that the interpolated activity should interfere in less degree with the retention of the original activity, where the conditions of association between the two activities are lessened. Will less retroactive inhibition occur where the conditions of association of the original activity with the interpolated activity are lessened? Is it possible to dissociate an interpolated activity from the original so that no retroactive inhibition will appear, in the relearning of the first, as a result of the learning of the second? Positive answers to these questions would be interpreted as evidence in favor of the Theory of Associative Interference.

The following experiments deal with four aspects of the general problem. If the first activity is mediated by vision and the second learning activity by audition, how will the retroactive effects compare with those resulting when both activities are mediated by vision? If the first activity takes place in one room environment and the second in a different room environment, how will the resulting retention compare with that of the control condition? If the original learning involves dextral manual activity and the second learning involves sinistral manual activity, how will retention compare with that in which both activities involve a common manual activity? If the original material is learned in the waking state and the interpolated material is learned in the hypnotic state, will retention be higher than if both materials are learned in the waking state?

Experiment I: Dissociation through Change in the Method of Presentation of Material

Problem.—If the interpolated material to be learned is presented visually and the original material is presented auditorially, to what degree will the interpolated material interfere with the retention of the original material?

Method.

The learning materials consisted throughout of lists of twelve nonsense syllables. These were printed in $\frac{3}{16}$" type on flexible cardboard. In the visible learning the material was presented by means of the Chicago memory drum. The syllables of a list were exposed at intervals of three seconds until learning was complete. The same procedure was followed during relearning. A four-second interval was interpolated between the successive presentations of the list.

The auditory material was presented under the same conditions, except that the syllables were spelled aloud (*e.g.*, F-I-H) by the experimenter. Care was taken that enunciation was clear. The order of the syllables of the lists was so arranged that rhyming of neighboring letters was eliminated, insofar as possible, in order that cues extraneous to the material should not influence the speed of learning.

The anticipation method of learning was used. The lists were presented until the subject could make one perfect written anticipation of the material. Interpolated learning began immediately after the completion of the learning of the first material. Recall and relearning followed twenty-four hours after the completion of the original learning.

The lists of syllables were circulated, throughout all of the conditions of the experiment, so as to equalize the effects of possible differences in difficulty, between any list or lists, upon any one condition or conditions. The different subjects learned the series of conditions in counterbalanced orders, in order to distribute the practice effects throughout the various conditions.

Twenty-four subjects, twenty men and four women, comprised the experimental group. All of the subjects, excepting two, were graduate students in psychology at the University of Chicago. The two untrained subjects were put through a preliminary training period in the learning of nonsense syllables before proceeding to the experimental work. All subjects worked in every condition of the experiment.

The conditions of the experiment were as follows:

	Method of Presentation of Material (Original)	Method of Presentation of Material (Interpolated)	Method of Presentation of Material at Recall
Condition A.....	Auditory	Auditory	Auditory
Condition B.....	Auditory	Visual	Auditory
Condition C.....	Visual	Auditory	Visual
Condition D.....	Auditory	Rest	Auditory
Condition E.....	Visual	Visual	Visual
Condition F.....	Visual	Rest	Visual

At the time of recall the following typical instructions were given: "Yesterday you learned two lists of nonsense syllables. The first list was learned visually and

the second list learned auditorially. I now want you to anticipate as many of the syllables of the first list as you can, on the first presentation of that material. The list will be presented to you in the same fashion as it was yesterday. Relearn this list as rapidly as possible."

Results.—Table I shows the effects of sensory dissociation of the interpolated material upon the retention of the first material learned. Data are presented in terms of the average savings and anticipation scores, the difference between these scores on the comparable conditions, and the reliability of these differences in terms of their probable errors.

TABLE I

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition A	.5370±.0275		.2099±.0300	
Condition B	.6397±.0152		.2732±.0245	
Diff.	.1027±.0271*	3.79	.0633±.0292	2.16
Condition E	.4941±.0311		.2033±.0198	
Condition C	.6113±.0181		.2999±.0264	
Diff.	.1172±.0335	3.49	.0966±.0305	3.16

* Since all of the subjects worked in all of the conditions of the experiment the probable error of the difference of the means formula involves the correlation of the data of the two conditions compared.

An examination of Table I reveals a greater savings score in Condition B, where the interpolated learning is sensorially dissociated from the original learning activity, than in Condition A, where there is no such sensory dissociation, but where the learning activities are mediated by a common sense avenue. The first lists were auditorially presented in both Condition A and in Condition B. The difference between the means of the savings scores is almost four times the probable error of the difference. This represents a satisfactory reliability. The difference between the average anticipation scores of the same two conditions is in the same direction as that between the savings score averages, but is not as reliable a difference.

In the same table, a comparison of the average savings scores of Conditions C and E, reveals that the dissociated

condition again shows reliably the higher score. The average anticipation score of Condition C is also greater than that of the control condition. The original learning in each of these cases was visual.

Table II shows the average savings and anticipation scores of the Conditions B and C, in each of which the second learning activities were sensorially dissociated from the first, in relation to those of Conditions D and F, respectively. In these latter conditions single lists only were learned.

TABLE II

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition D.....	.6807±.0226		.5032±.0383	
Condition B.....	.6397±.0152		.2732±.0245	
Diff.....	.0410±.0199	2.06	.2300±.0419	5.48
Condition F.....	.6472±.0190		.4057±.0317	
Condition C.....	.6113±.0181		.2999±.0264	
Diff.....	.0359±.0250	1.43	.1058±.0401	2.63

It will be noted in Table II that the average savings score of Condition B, where the first activity was mediated by audition, and the second by vision, approaches that of Condition D, in which a single list was auditorially learned. The same situation holds in the comparison of Conditions F and C. The differences in anticipation scores appear to be more reliable than those in savings scores in both comparisons.

In Table III comparisons are made of the average savings

TABLE III

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Conditions A+E.....	.5155±.0210		.2066±.0179	
Conditions B+C.....	.6255±.0119		.2865±.0175	
Diff.....	.1100±.0242	4.54	.0799±.0251	3.18
Conditions F+D.....	.6640±.0162		.4545±.0262	
Conditions B+C.....	.6255±.0119		.2865±.0175	
Diff.....	.0385±.0201	1.91	.1680±.0315	5.33

and anticipation scores of the control Conditions B and C combined, with those of the dissociated Conditions A and E combined, and with those of the single list Conditions F and D combined.

When the experimental Conditions B and C are combined and compared with the combined control Conditions A and E, the differences in average savings and anticipation scores become even more reliable. A combination of the average savings and anticipation scores of the single lists F and D, when compared with a corresponding combination of the experimental Conditions B and C, shows a reliable difference in average anticipation scores, but not in average savings scores.

Reed (**20**) has shown that if the material to be learned is presented auditorially or visually, retention is higher when the relearning proceeds by the same method of presentation as that employed during learning.

The results of Experiment I reveal that less retroactive inhibition results when the first and the second activities separately involve the auditory or the visual sensory tracts than when the two activities involve the same sensory pathways. This result may be interpreted in terms of the Theory of Associative Interference, which assumes that retroactive inhibition increases as the number of common neural elements involved in the two activities increases.

Summary.—1. When two activities are learned, and the first of these is mediated by vision and the second by audition, or vice versa, there is less retroactive inhibition than when both activities involve the same visual or auditory pathways.

2. When two learning activities are sensorially dissociated the amount of retroaction of the second material on the retention of the first, as measured by the savings score, is not reliably higher than that which occurs when no interpolated activity is introduced. However, if the differences in anticipation scores are considered, there are reliable differences in favor of the single list conditions.

Experiment II: Dissociation through Change of Environment

Problem.—Retroactive inhibition may possibly be traced to the association of the first and the interpolated materials to be learned with the extra-organic setting in which they occur. Thus the same stimulating conditions may bring about an interference between the two activities by cross association. If the original and the second learning activities occur in one environment, which is common to both, then, since both would tend to be associated with the same environment, there should be a number of associations formed between the two activities and large interference should result.

The problem of this experiment is to determine whether a lesser retroactive inhibition follows when the first activity is learned in one environment and the interpolated activity is learned in a second environment, than that which follows the control condition, in which both activities, original and interpolated, occur in the same environment.

Method.

The method of learning was essentially the same as that outlined under Experiment I except that all learning materials were presented visually. Where a second list was to be learned, one minute elapsed between the completion of the learning of the first and the beginning of the learning of the second. This interval was allowed in order that the subject could pass to the second room for the interpolated learning, in the condition where this was necessary.

Twenty-four subjects, all students in psychology at the University of Chicago, comprised the experimental group. The three undergraduate students in the group were unpracticed in the learning of nonsense syllables and were trained, therefore, in this function for three periods preliminary to the beginning of the experimental work. Nineteen men and five women served as subjects. All subjects worked in each condition of the experiment.

The conditions of the experiment were as follows:

	Learning Environment		Relearning Environment
	List I	List II	List I
Condition A.....	Room I	Room I	Room I
Condition B.....	Room I	Room II	Room I
Condition C.....	Room I	Rest	Room I

It was attempted to make the room environments in which the learnings occurred strikingly different from each other. The rooms differed in size and in the nature of

the arrangements of their contents. In each the experimenter oriented himself in space with respect to the subject in a different manner. The memory drums used were different models of the Chicago drum. Room I was lighted by a Mazda lamp which brought all of the environment into clear relief. In Room II the illumination consisted of a beam of red light which issued from an aperture cut in a tin hood, which otherwise completely enclosed the source of illumination. The beam of light was directed upon the face of the drum on which the learning material was exposed. Sufficient light was deflected so that the subject could write his anticipation of the material without being handicapped in the task by insufficient illumination. The room environment, beyond the memory drum, was thus poorly lighted. Pressey (**19**) has suggested a relationship between the speed of learning and the degree of illumination. Three students judged the illumination of the materials to be learned in the two environments to be equal.

Before the relearning began the subjects were given the following typical instructions: "Yesterday you learned two lists of nonsense syllables. The first list was learned in this room and the second list in the other room. I now want you to anticipate as many syllables of the first list as you can on the first presentation of that material. Relearn this list as rapidly as possible."

Results.—Table IV shows the average savings and anticipation scores of the various conditions of the experiment, the differences between these, and their probable errors.

TABLE IV

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition B	.4680±.0316		.1770±.0238	
Condition A	.4526±.0319		.1874±.0236	
Diff.	.0154±.0419	0.37	.0104±.0294	0.35
Condition C	.6790±.0209		.4096±.0305	

An examination of the average scores presented shows no noteworthy differences between the respective averages of the experimental Condition B and its control Condition A.

There have been many experimental findings which indicate that a change at the time of recall in the sensory environment in which a habit is fixated has a detrimental effect on that recall. One possible explanation of this detrimental effect would be the absence of the proper stimulating conditions.

Yum (**29**) has demonstrated that stimuli, similar to those which were originally associated with an activity, may call

forth that activity. Stimuli which approach each other in similarity tend to call forth the same response.

In Experiment II we wished to determine whether there was a greater or a lesser retroactive inhibition when opportunity was given for both the interpolated and the original activities to be associated with the same environment rather than with different environments. From the past work on forgetting it would seem probable that a first and a second activity, each associated with different stimulating conditions, would lead to as great a degree of retroactive inhibition and interference as when these two activities were associated with a common environment. The results as presented in Table IV show no reliable differences resulting from a dissociation of the environments.

Summary.—1. When the original learning activity occurs in one environment and the interpolated activity occurs in a second environment, the resulting savings and anticipation scores are no greater than when both learning activities occur in a common environment.

2. Since no more retroactive inhibition is present in the condition where both activities occur in different environments than is to be found in the condition where the environment is common to the two activities, it appears that the presence of a common room environment does not mediate a closer association between the original and the interpolated activities.

Experiment III: Dissociation through Change in Manual Activity during Learning

Problem.—The problem that is the concern of Experiment III is the following one. If the original activity involves the left hand and the interpolated activity involves the right hand, to what degree will the interpolated activity interfere with the retention of the first activity?

Webb (**25**) studied transfer of training and retroactive inhibition in rats and human subjects. The experimental data were gathered from maze learning activities. Among the problems he attacked were: (*a*) the dependence of transfer

upon the nature of the first and the second learning activities; and (*b*) the possible retroactive inhibition effected by one or more activities upon the other. He reports that (1) there were practically no cases of habit interference when total errors, time, and trials were the criteria considered; and (2) the degree of similarity of the mazes and the nature of the first and the second activities were the factors found to influence the degree of transfer. Wiltbank (**27**) found, while training rats in maze habits, that a positive transfer in both time and errors occurred from one maze to another. Brockbank (1) trained rats on a circular maze until mastery. A thirty-day training period then followed on a rope-ladder problem. Retention of the first habit was then tested. The interpolated activity did not appear to decrease the average retention. When the interpolated activity consisted of training on a saw-dust problem box, the same general lack of interference with the retention of the first habit was established.

Method.

The method by which learning proceeded was essentially the same as that of Experiment II. The exposure time was four seconds per syllable since this period was found to be the optimum for left-handed writing. Interpolated learning followed immediately after the completion of the original learning. The subject wrote the material as it was learned, with either the right or the left hand, according to the condition of the experiment.

The experimental group was composed of twenty-four subjects, four women and twenty men. All of the subjects were graduate students in the University of Chicago. All subjects worked in every condition of the experiment.

The conditions of the experiment were as follows:

	Hand of Motor Activity		
	List I	List II	Recall List I
Condition A.....	Right	Left	Right
Condition B.....	Left	Right	Left
Condition C.....	Left	Rest	Left
Condition D.....	Left	Left	Left
Condition E.....	Right	Rest	Right
Condition F.....	Right	Right	Right

At the time of relearning the subjects were instructed typically as follows: "Yesterday you learned two lists of nonsense syllables. You wrote the first list with the

left hand as you learned it. The second list was written with the right hand. Relearn the first list as soon as possible, writing with your left hand."

Results.—Table V shows the average savings and anticipation scores of those conditions of the experiment in which the first and the second activities involved different hands, compared with the corresponding scores of those conditions in which the activities involved the same hands. Differences are shown between the average scores and their probable errors are indicated.

An examination of Table V shows that when both the first and the second activities involve different hands there

TABLE V

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition A	.5336±.0237		.2246±.0187	
Condition F	.5048±.0328		.1992±.0212	
Diff.	.0288±.0330	0.87	.0254±.0253	1.00
Condition B	.5829±.0323		.2934±.0308	
Condition D	.5099±.0303		.2065±.0286	
Diff.	.0730±.0409	1.78	.0869±.0270	3.22

is a tendency toward greater savings and anticipation scores than when both habits involve the same hand. The reliability of the differences is not high but the differences are consistent in the direction of their trends.

Table VI shows the difference in average savings and

TABLE VI

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition E	.6472±.0190		.4057±.0317	
Condition A	.5336±.0237		.2246±.0187	
Diff.	.1136±.0317	3.58	.1811±.0360	5.03
Condition C	.7059±.0163		.4637±.0263	
Condition B	.5829±.0323		.2934±.0308	
Diff.	.1230±.0379	3.24	.1703±.0358	4.81

anticipation scores between Conditions A and E. In the first condition the original activity involved the right hand and the interpolated activity the left hand. Condition E involved a single right hand activity. Likewise, the average retention scores of the dissociated Condition B and the single activity Condition C are compared. In these latter two conditions the first activity involved the left hand.

The table shows highly reliable differences between the dissociated conditions and the single list conditions. This finding is to be contrasted with that of the experiment on sensory dissociation, where no reliable difference was found between the average savings scores of the dissociated lists and those of the single list conditions.

Table VII shows the differences between the combined savings and anticipation score averages of the dissociated Conditions A and B and those of the combined control Conditions D and F.

TABLE VII

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Conditions A+B.....	.5582±.0202		.2590±.0177	
Conditions D+F.....	.5073±.0225		.2028±.0182	
Diff.....	.0509±.0258	1.97	.0562±.0234	2.40

The results indicate that when the first of two activities involves a motor process different from that of the second there is a tendency toward less retroactive inhibition than when the original and the interpolated activities involve the same motor process. While the differences are in no case significantly reliable the general unbroken trend of the data is a factor to be considered.

Summary.—When the interpolated activity involves a different manual motor process from that of the first activity there is a consistent tendency towards a higher savings and anticipation score than when both activities involve the same motor pathways.

Experiment IV: Dissociation through Hypnosis

Problem.—The literature of abnormal psychology frequently makes reference to certain phenomena, such as double personality, automatic writing, hypnosis, etc. Many experiences present in these states seem to have no connection with those of the waking state, since they appear to be unrecallable when the subject is in a normal state. A noteworthy phenomenon of the hypnotic state is a frequently appearing amnesia for its events when the subject is awakened. There appears to be a dissociation of the experiences of the two states.

Such a dissociation, it is expected, would produce a lower degree of associative interference when the original and the interpolated materials are learned in the hypnotic and the waking states, respectively, than that produced when both are learned in the hypnotic state. Such a finding could be interpreted as further evidence in favor of the Theory of Associative Interference.

It is the purpose of Experiment IV to determine whether or not less retroactive inhibition follows the learning of a second material in a state dissociated from the first. The problem may be stated more explicitly as follows: If the interpolated material is learned while the subject is in the hypnotic state and the original material is learned in the waking state, to what degree will the interpolated material interfere with the retention of the original material?

Method.

The learning tasks consisted in the learning of nonsense material to the point of one perfect repetition. The lists were eleven syllables in length. These were presented by means of the Hull memory apparatus-at the rate of one syllable each two seconds. Four seconds elapsed between the successive presentations of a list. The subjects anticipated the material by spelling it aloud.

When an interpolated material was involved in a condition, learning commenced ten minutes after the completion of the first learning. This period of time was allowed for the induction of the hypnotic state. Thus, this period of time elapsed in all of the conditions of the experiment before the interpolated learning began. Relearning followed fifty-five minutes after the completion of the original learning and proceeded by the same anticipation method. During the rest periods the subjects were engaged in the reading of light narrative material. After the completion of the interpolated learning, the subjects always spent the interval of rest in the waking state before relearning.

Ten male subjects were used in the experiments. All subjects went through all of the conditions of the experiment, but in different orders, so as to distribute the effects of practice throughout them. Unpracticed subjects were given three days of preliminary practice in nonsense syllable learning technique. The various syllable lists were circulated throughout the conditions of the experiment, insofar as possible, so that difficulty of list might not influence the comparative average scores. Only volunteer subjects were used, none of whom were reimbursed for their work.

The method of inducing the hypnotic state was primarily that outlined by McDougall (**9**). Various authors have advanced different criteria of the presence of the hypnotic state. Hull (**7**) gives an excellent discussion and criticism of these criteria. He concludes that probably the only psychological criterion of hypnosis is a heightened susceptibility to suggestion. Certain phenomena, however, usually appear in a deep trance state. Some of these are waxy plasticity, post-hypnotic amnesia, analgesia, negative and positive hallucinations. All or some of these phenomena may be present. However, a trained observer can usually detect the presence of the somnambulistic state by the total picture of the behavior of the hypnotized subject.

Only those subjects showing the deeper phenomena of the hypnotic state were used in the experiments here reported. The many subjects who appeared incapable of the phenomena of deep hypnosis were discarded. After the experimental work began no suggestions of post-hypnotic amnesia were made to the subjects for fear that the suggestion might lead them to simulate amnesia when there was no true amnesia present. This procedure is in contrast to that of Mitchell (**16**) who always gave suggestions of post-hypnotic amnesia to her subjects during the hypnotic trance.

The conditions of the experiments were as follows:

	Condition of Learning		Condition of Relearning
	Original Material	Interpolated Material	Original Material
	Experiment 1		
Condition A.....	Waking	Waking	Waking
Condition B.....	Waking	Trance	Waking
Condition C.....	Waking	Trance Rest	Waking
Condition D.....	Waking	Waking Rest	Waking
	Experiment 2		
Condition A.....	Trance	Trance	Trance
Condition B.....	Trance	Waking	Trance
Condition C.....	Trance	Waking Rest	Trance
Condition D.....	Trance	Trance Rest	Trance

Results.—Table VIII shows the average savings and anticipation scores of the conditions of Experiment I and the differences between the means of the experimental and their control conditions.

TABLE VIII

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition A	.3795±.0296		.1726±.0227	
Condition B	.4965±.0228		.2545±.0294	
Diff.	.1170±.0293	3.99	.0819±.0286	2.86
Condition C	.6265±.0381		.5545±.0421	
Condition D	.4877±.0317		.4273±.0318	
Diff.	.1388±.0396	3.51	.1272±.0420	3.03

An examination of the table shows a higher average anticipation and savings score for the dissociated conditions. While some of the critical ratios fall below the generally accepted standard of reliability, all of the differences are in a uniform direction. It should be noted how closely the average savings score of the dissociated Condition B approaches that of the single list Condition D. However, the average anticipation scores of the two last conditions mentioned are apparently reliably different.

Table IX shows the average savings and anticipation scores

TABLE IX

	Average Savings Score	D./P.E.D.	Average Anticipation Score	D./P.E.D.
Condition A	.4837±.0308		.2272±.0303	
Condition B	.6375±.0265		.3182±.0380	
Diff.	.1538±.0359	4.28	.0910±.0403	2.26
Condition C	.7827±.0174		.5908±.0448	
Condition D	.6429±.0294		.4908±.0288	
Diff.	.1398±.0318	4.38	.1000±.0437	2.29

of the conditions of Experiment II and the differences between the means of the experimental and their control conditions.

The directions of the differences between the means of the comparable conditions of Experiment II are uniform. The dissociated conditions again show an advantage over the

control conditions in average savings and anticipation scores. While the savings score differences appear to be true differences, the anticipation scores are not as reliable with regard to their differences.

It is to be noted that a higher average retention is to be found in the conditions of Experiment II. It thus appears that retention is better when recall and relearning occur in the hypnotic state than when they both occur in the waking state. This finding is in accord with that of the majority of the findings on the subject, as typified by the results of the experiments of Young (**28**), Dallenbach and Jenkins (**4**), etc.

The differences between the findings of this experiment and those of Mitchell (**16**) may possibly be traced to the differences in technique employed. Mitchell worked with but two subjects and, hence, considered her work largely exploratory in nature and her findings as tentative. She gave suggestions of post-hypnotic amnesia in those conditions in which hypnosis was employed. What effects such suggestions would have upon the average retention is not known. It is apparent from our study that post-hypnotic amnesia may result in some cases without suggestions to that effect. As far as can be determined from the description of her method, retention tests followed immediately after the interpolated activity. Further, she allowed a two-minute interval for the induction of the hypnotic trance whereas a ten-minute interval was allowed in these experiments.

The work in the field of hypnotic phenomena is still at an exploratory level. Authors are still in disagreement as to what constitutes hypnosis. Some regard somnambulance, only, as hypnosis, while others broaden the content of the term to include even the lightest phenomena.

It appears from the results here reported that when a second activity is dissociated from a first activity, by means of hypnosis, the resulting retention is higher than when no such a dissociation occurs. The dissociation of the two activities appears to be more or less complete, as evidenced by the proximity of the average savings scores of the dis-

sociated conditions to those of the comparable single list conditions.

Summary.—1. When the interpolated activity occurs in the hypnotic state and the original activity in the waking state, the resulting anticipation and savings scores are higher than when both activities occur in the waking state. The converse is also true.

2. Materials learned in the dissociated conditions show average savings scores closely approaching in size those of the comparable single list conditions.

3. When material is learned in the hypnotic state a greater retention is apparent than when material is learned in the waking state.

General Summary and Conclusions

The experiments of the present study were undertaken in an attempt to determine to what extent the amount of retroactive inhibition, following interpolated activity, would be reduced by lessening the possibility of associative connection between the original and the interpolated activities.

Four types of dissociation were studied: those involving (*a*) sensory mode of presentation of the material; (*b*) room environment; (*c*) hypnotic and waking states; (*d*) different motor activities.

The dissociation of the sensory modes of presentation of the materials had a significant effect upon the amounts retained. When the sensory mode of presenting the second materials was changed from that of the first, not only did higher average anticipation and savings scores result, but the average savings scores closely approached those of the conditions in which no interpolated material was introduced.

The dissociation of the two room environments in which the first and the second learning activities occurred led to approximately the same amount of retroactive inhibition as occurred in that condition where the two activities occurred in the same room environment.

When the first of two learning activities involved the motor pathways of one hand and the second the motor path-

ways of the other hand, there appeared to be less resultant retroactive inhibition than when both activities involved the same motor elements. The differences in retention scores, while not highly reliable, were uniform in direction and, at least, were indicative of a tendency.

The dissociation of the original and the interpolated activities, through the hypnotic technique, led to less retroactive inhibition than occurred where the two activities proceeded in the same state. The dissociation of the two activities gave average retention scores which closely approached those of the comparable condition in which but a single list was learned.

The experimental results here reported, except those on the dissociation of room environments, may be interpreted to lend further support to the Theory of Associative Interference as one of the explanations of retroactive inhibition.

(Manuscript received October 3, 1934)

References

1. Brockbank, T. W., Redintegration in the albino rat. *Animal Beh. Mono.*, 1919, **4**, 1–65.
2. Bunch, M. E., and Wientge, K., The relative susceptibility of pleasant, unpleasant, and indifferent material to retroactive inhibition. *J. Gen. Psychol.*, 1933, **9**, 157–78.
3. Cheng, H. Y., Retroactive effect and degree of similarity. J. Exper. Psychol., 1929, **12**, 444–49.
4. Dallenbach, J. M., and Jenkins, J. C., Obliviscence during sleep and waking. *Amer. J. Psychol.*, 1924, **35**, 606–12.
5. De Camp, J. E., A study of retroactive inhibition. *Psychol. Mono.*, 1915, **19**, No. 84, 1–69.
6. Harden, L. M., A quantitative study of the similarity factor in retroactive inhibition. *J. Gen. Psychol.*, 1929, **2**, 421–30.
7. Hull, C. L., *Hypnosis and suggestibility: an experimental approach.* New York: Appleton-Century, 1933.
8. Johnson, L. M., Similarity of meaning as a factor in retroactive inhibition. *J. Gen. Psychol.*, 1933, **9**, 377–89.
9. McDougall, W., *Outline of abnormal psychology.* New York: Scribner's, 1926.
10. McGeoch, J. A., Studies in retroactive inhibition: I. The temporal course of the inhibitory effects of interpolated learning. *J. Gen. Psychol.*, 1933, **9**, 24–45.
11. McGeoch, J. A., Studies in retroactive inhibition: II. Relationships between temporal point of interpolation, length of interval, and amount of retroactive inhibition. *J. Gen. Psychol.*, 1933, **9**, 44–57.

12. McGeoch, J. A., The influence of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, 1932, **44**, 695–708.
13. McGeoch, J. A., and McDonald, W. T., Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, **43**, 579–88.
14. McGeoch, J. A., and McKinney, F., Retroactive inhibition in the learning of poetry. *Amer. J. Psychol.*, 1934, **46**, 19–33.
15. McGeoch, J. A., and Nolen, M. E., Studies in retroactive inhibition: Temporal point of interpolation and degree of retroactive inhibition. *J. Compar. Psychol.*, 1933, **15**, 407–17.
16. Mitchell, M. B., Retroactive inhibition and hypnosis. *J. Gen. Psychol.*, 1932, **7**, 342–59.
17. Müller, G. E., and Pilzecker, A., Experimentelle Beiträge zur Lehre vom Gedächtniss. *Zeits. f. Psychol.*, 1900, Ergs. 1, 1–300.
18. Peters, H. N., *Mediate association.* Unpublished Doctor's Dissertation, Department of Psychology, University of Chicago, 1933.
19. Pressey, S. L., The influence of color upon mental and motor efficiency. *Amer. J. Psychol.*, 1921, **32**, 326–56.
20. Reed, H. J., *The influence of a change of conditions upon the amount recalled.* Unpublished Doctor's Dissertation, Department of Psychology, University of Chicago, 1930.
21. Robinson, E. S., Some factors determining the degree of retroactive inhibition. *Psychol. Mono.*, 1920, **34**, No. 128, 1–56.
22. Robinson, E. S., and Heron, W. T., Results of variations in length of memorized materials. J. Exper. Psychol., 1922, **3**, 428–48.
23. Skaggs, E. B., Further studies in retroactive inhibition. *Psychol. Mono.*, 1925, **34**, No. 161, 1–60.
24. Tolman, E. C., Retroactive inhibition as affected by conditions of learning. *Psychol. Mono.*, 1917–18, **25**, 1–50.
25. Webb, L. W., Transfer of training and retroaction. *Psychol. Mono.*, 1917, **24**, 1–90.
26. Whitely, P. L., The dependence of learning and recall upon prior intellectual activities. J. Exper. Psychol., 1924, **7**, 420–28.
27. Wiltbank, R. T., Transfer of training in white rats upon various series of mazes. *Animal Beh. Mono.*, 1919, **4**, 1–65.
28. Young, P. C., An experimental study of mental and physical functions in the normal and hypnotic states. *Amer. J. Psychol.*, 1925, **36**, 214–32.
29. Yum, K. S., An experimental test of the law of assimilation. J. Exper. Psychol. 1931, **14**, 68–82.

A STUDY OF LATENT LEARNING

BY MERRILL T. EATON

Indiana University

This study proposes to investigate the 'latent learning' resulting from practice in drawing circles without knowledge of results; and to measure the objective learning resulting from such practice. The problem, therefore, is a dual one.

The term 'latent learning' is borrowed from Blodgett's (**2**) study of the effect of the introduction of reward upon the maze learning of rats. Blodgett found that a group of rats which ran the maze six days in succession, but were not rewarded with food at the end of each trial, did not improve in maze performance. That skills, however, were accumulating during these six trials without food reward was shown by a great drop in the number of errors on the first trial following the day on which the rats found food in the food box. Blodgett has named this learning, which does not show objectively until there is practice with reward, 'latent learning.' The findings of Tolman and Homzik (**12**) in a similar investigation agree with those of Blodgett.

The term 'objective learning' as used in this investigation refers to improvement in performance which results directly from practice without knowledge of results and which may be measured without consideration of latent learning.

The problem of learning through practice without knowledge of results has been attacked from a number of different angles with somewhat conflicting results. One of the early studies dealing specifically with the problem of learning through practice without knowledge of results was that of Judd (**7**). The problem presented in Judd's investigation was to determine the correct continuation of lines when only a part of the line was exposed to view and the remainder hidden by a screen. The lines were drawn from the center of the paper to the right-hand edge at nine different angles to the

horizontal. When one of these lines was presented, the subject was required to mark with a pencil on the screen what he regarded as the exact continuation of the line. The nine lines in the series were presented 20 times each day for a period of 10 days. Judd concluded from the results of this study that learning does not take place as the result of mere practice.

Spencer (**10**) criticized Judd's method of adding algebraically the subjects' plus and minus errors in calculating their average errors. By disregarding signs Spencer found an improvement of 11 per cent in performance. Repeating Judd's experiment and computing the errors without regard to signs, Spencer found that three of his four subjects improved (learned) even though they had no knowledge of the accuracy of their responses.

Another extensive investigation concerning the question of whether there is learning through practice without knowledge of results was made by Thorndike (**11**). Thorndike presented the problem of drawing lines four inches long. The subjects were not informed as to the accuracy of their responses and there was no improvement in performance. Smith (**9**), however, in a series of experiments in learning to judge the length of lines without knowledge of results found that mere repetition was an effective factor in learning.

The problem of learning to judge the length of strips of cardboard without knowledge of results was studied by Hall (**6**). Hall divided his subjects into five experimental groups, one of which made judgments without a standard of known length being present, and the other four groups with one or more standards of known lengths. The group making judgments without the presence of a known standard did not improve. In the other groups, however, there was improvement; and Hall concludes improvement may take place in length-judging even though the subjects are not informed of the accuracy of their responses, provided a standard of known length is present.

Numerous studies, of which those of Book and Norvell (**3**), Ross, (**8**) Crafts and Gilbert (**4**), and Arps (**1**) are typical, are

somewhat concerned with the problem of learning through mere repetition. Their chief concern, however, seems to be the measurement of the effect on learning of certain incentives.

None of the studies described has attempted to measure the latent learning resulting from practice without knowledge of results, with the possible exception of Blodgett and Tolman and Homzik. Blodgett's technique, however, offers an interesting approach to a study of both objective and latent learning resulting from practice without knowledge of results.

The lack of agreement in the findings of studies of learning through practice without knowledge of results may be explained on the basis that in some learning situations practice without knowledge of results will immediately show objectively in improvement of performance while in other types of learning improvement will not show until there is practice with knowledge of results, although skills may have been accumulating during the practice periods. The sense or senses involved and the type of activity predominating, motor or ideational, may also constitute important factors in the problem. Whether the investigation studies the acquisition of new responses or whether the study is concerned with the improvement of an already learned activity is also a question to be considered.

Experimental Conditions

The purpose of this study is to measure both the latent learning and the objective learning resulting from practice in drawing circles without knowledge of results, and to compare and contrast the extent of each type of learning. To make these measurements accurate and objective the following plan was followed:

1. A new apparatus (rotameter) was devised which measured each trial circle accurately from which the measurement on any one trial is directly comparable to any other trial, either for the same subject or between subjects.

2. Sixteen subjects were tested on the rotameter.

3. The subject was never permitted to see the rotameter or his record.

4. Each subject was given 100 trials on the rotameter and his average error computed.

5. The sixteen subjects were divided into two groups, experimental and control, each of which contained eight subjects selected on the basis of their average errors on the 100 trials. The subjects in each group were paired so that the individual subjects were nearly identical in initial performance and so that the average error for the two groups was approximately the same.

6. The subjects comprising the experimental group were given 1000 practice trials without knowledge of results.

7. All subjects were given 100 practice trials with knowledge of results.

8. All subjects were also given another 100 practice trials with knowledge of results.

Apparatus.—The apparatus devised to measure learning in this experiment is a rotameter. In Fig. 1 the apparatus is shown. The base (*A*) is of hardwood, 1″ × 8″

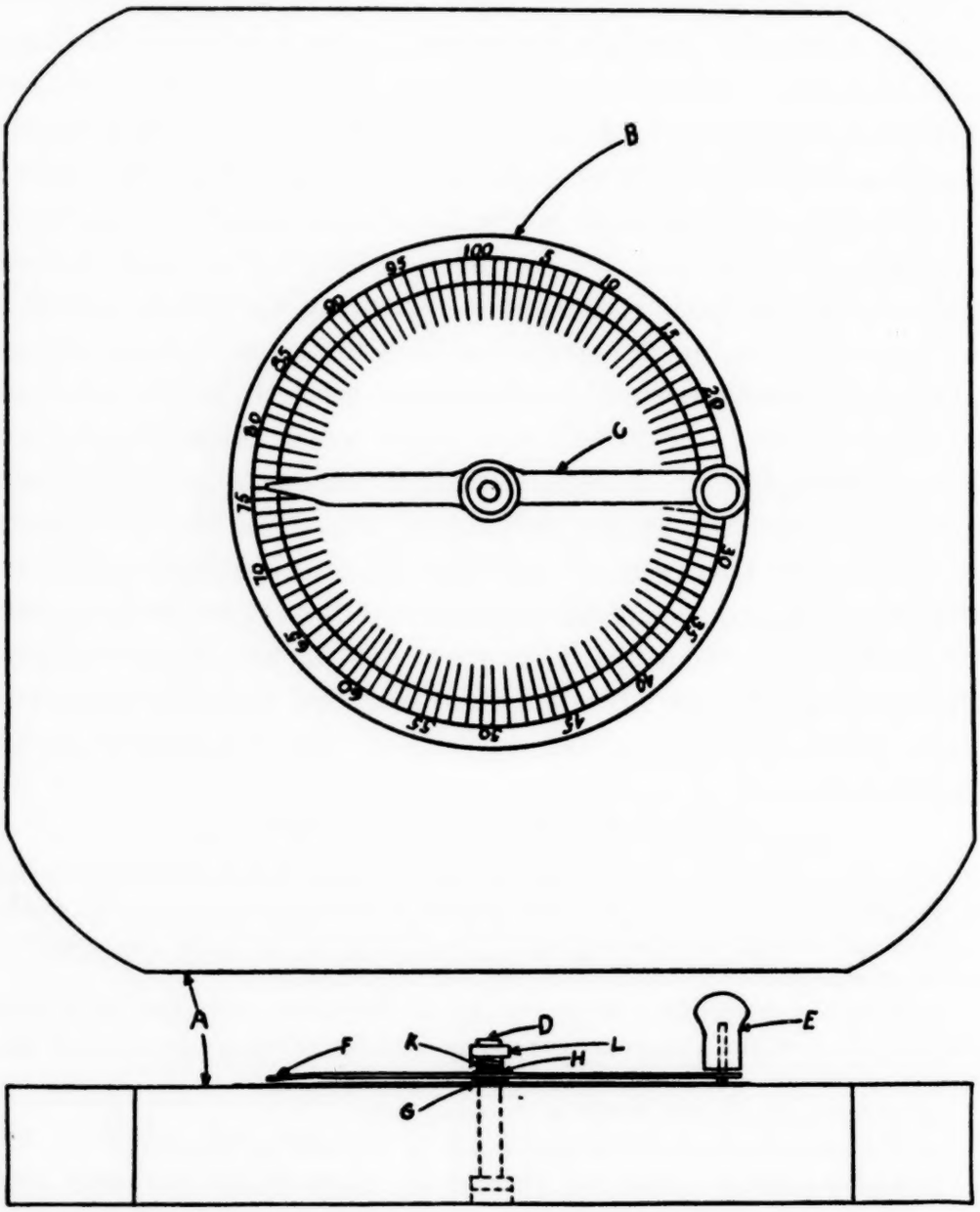

FIG. 1. The rotameter.

× 8″. The dial (*B*) is a graduated disk divided into 100 degrees. The metal needle (*C*) swings on the shaft (*D*) which protrudes through the base and graduated disk. The needle has a pencil-shaped stylus (*E*) on one end, and a pointer (*F*) on the other. The stylus is metal and turns free on its fitting into the needle. The center of the stylus handle is 2″ from the center of the shaft. The pointer is turned down slightly so that

fine discriminations can be made. There is a washer (*G*) under the needle to raise it slightly above the graduated disk so that it swings clear. Above the needle there is another washer (*H*), a coil spring (*K*) and lock nut (*L*). By the use of the lock nut and spring any desired tension can be placed on the needle and, once adjusted, remains the same. The objectivity, accuracy, and ease of operation of the apparatus can be readily seen. All circles drawn have exactly the same radius regardless of the amount of under or overlapping.

Subjects.—Sixteen college students were used as subjects. The age, sex and class standing of these subjects are given in Table I.

TABLE I

AGE, SEX, AND CLASS DISTRIBUTION OF SUBJECTS

Age					Sex		Class	
18	19	20	21	22	Male	Female	Freshman	Sophomore
6	5	3	1	1	6	10	11	5

Practice without Knowledge of Results.—All testing was done in a quiet laboratory where distractions were held to a minimum. Only the experimenter[1] and one subject were in the laboratory during a testing series. When the subject entered the room he was seated at a low table and blindfolded. The hoodwink was a large bag, made of heavy, black sateen, with draw strings at the bottom. The bag was placed over the subject's head and the draw strings were tightened snugly around the neck excluding any possibility of seeing. Slits were cut in the top and back of the bag for ventilation The rotameter was then removed from a drawer in the desk and placed before the subject. *The subjects were never permitted to see the apparatus, nor their scores. Neither were they informed of the radius of the circles they were drawing.* The instructions to the subjects were as follows:

"I am interested in finding out how accurately you can learn to draw circles without knowing the accuracy of your responses. The apparatus you are to use is so constructed that all circles you draw will have the same radius. I should like to have you do your very best on each trial."

The subject's hand was placed on the stylus and he was asked to rotate the stylus until he considered he had drawn a circle. Then his hand was removed. The experimenter then gave the needle a spin to assure an unselected starting point, and the subject's hand was again placed on the stylus and the procedure repeated. Before beginning the recording of data, each subject drew three practice circles to familiarize himself with the apparatus.

In recording the data, the experimenter noted the starting and finishing position of the needle point on the scale. The difference between these two points in degrees constituted the error. Each degree was approximately one-eighth inch, $(4 \times 3.1416)/100$. If the circle was not completed, the error was given a minus sign; if the circle overlapped, the error was given a plus sign. During the drawing of the circle, there was no contact of the subject's arm, hand, or fingers between the shoulder

[1] Babb, V. W., a graduate student, assisted in collecting the data.

and where the fingers grasped the stylus. A *trial series* consisted of 50 trials; 25 trials given in succession, a rest period of five minutes, and then the remaining 25 trials. When the subject finished his trial series, the rotameter and data sheets were removed from view before the blindfold was lifted.

Practice with Knowledge of Results.—The testing procedure for the trials with knowledge of results was exactly the same as that for the trials without knowledge of results, except that immediately following each trial the subject was informed as to the extent of his error. This information was given in eighths of an inch. Before starting a practice series, each subject was shown on a ruler one-eighth of an inch. If the subject failed by two degrees to close the circle, he was told, "You failed by two-eighths of an inch to close the circle." If he overlapped two degrees, he was told, "You overlapped two-eighths of an inch." Before beginning the 'with knowledge of results' trials, the meaning of the terms 'failure to close' and 'overlap' were demonstrated to the subject by the experimenter's drawing pencil circles and indicating those which did not close and those which overlapped.

Practice Situation.—There were four different practice situations:

PS1. The first 100 trials without knowledge of results. (All 16 subjects.)
PS2. The 1,000 practice trials. (Only the 8 subjects of the experimental group.)
PS3. The first 100 trials with knowledge of results. (All subjects.)
PS4. The second 100 trials with knowledge of results. (All subjects.)

Pairing of Subjects.—The 16 subjects were given 100 trials on the rotameter and paired on the basis of their average errors. These data are given in Table II. As

TABLE II

THE AVERAGE ERRORS MADE BY THE SUBJECTS OF THE EXPERIMENTAL AND THE THE CONTROL GROUPS ON THEIR FIRST 100 TRIALS

Pairs	Experimental Group		Control Group		Diff.
	Subject	Aver. Errors	Subject	Aver. Errors	
1– 9	1	3.35	9	3.21	−.14
2–10	2	3.61	10	4.03	+.42
3–11	3	7.67	11	7.89	+.22
4–12	4	5.84	12	5.78	−.06
5–13	5	6.33	13	6.67	+.34
6–14	6	8.59	14	8.03	−.56
7–15	7	4.25	15	4.10	−.15
8–16	8	4.17	16	4.41	+.24
Average		5.476		5.515	.039

shown in Table II, subjects were paired whose average errors were similar. Thus, subjects 1 and 9 were paired. Subject 1 had an average error of 3.35 and subject 9, 3.21, a difference of .14 error (an error is one degree ⅛ inch). Following the plan of pairing those with similar errors, the two groups, subjects 1, 2, 3, 4, 5, 6, 7, and 8, and subjects 9, 10, 11, 12, 13, 14, 15, and 16 were formed. The first named subjects were designated the experimental group and the latter subjects the control group. The largest difference between the average error of any pair is .56 (pair 6–14) and the smallest is .06 (pair 4–12). The difference between the average errors of the two groups, how-

ever, is only .039 error which has a statistical reliability of only slightly more than chance (Table X). The two groups were, therefore, of practically equal ability at the beginning of the experiment.

RESULTS

Experimental Group.—Following the 100 initial trials, the subjects of the experimental group were given 1,200 trials (in series of 50), PS2, 1,000 trials without knowledge of results, PS3, 100 trials with knowledge of results, and PS4, 100 trials with knowledge of results. These data are given in Table III. In addition to the average error made by each subject for each practice situation, PS1 performance is compared with that of PS2, PS3, PS4, and PS3 plus PS4. Improvement is shown both in the reduction of errors and gain in percentages. Reading from top to bottom the performance of the individual subjects for any practice situation may be examined. Reading from left to right, the reduction in errors and gain in percentages for each practice situation may be studied. The bottom row of figures is group averages, group error reduction, and group gain in percentages. For PS2, individual reduction in errors varied from .28 errors to 5.23 errors which transferred into percentages range from 6.7 to 60.9 percent gain in efficiency. Individual variations in error reduction and percentage gain for the other practice situations may be determined by examining the other columns of Table III.

The group average error reduction in PS2 was 2.14 errors which constituted a 39.1 percent gain over PS1. For PS3, the average group reduction in errors over PS1 was 3.28 errors, a 60 percent improvement. The group reduction in errors for PS3 over PS2 was 1.14 errors or 34.1 percent. This gain is only slightly less than that accomplished by the 1,000 practice trials, 34 percent gain compared with 39.1 percent. Only a 5.9 percent gain was achieved on PS4 over PS3. This gain is quite insignificant when compared with that of PS2 over PS1 and PS3 over PS2. If the average error for both the 100 trials with knowledge of results is compared with initial performance there is a reduction of 3.35 errors which is an improvement of 61.1 percent.

The reliability of the differences between the group aver-

TABLE III

THE AVERAGE ERRORS MADE BY THE SUBJECTS OF THE EXPERIMENTAL GROUP ON THEIR FIRST 100 TRIALS COMPARED WITH THEIR 1,000 PRACTICE TRIALS AND THE 200 TRIALS WITH KNOWLEDGE OF RESULTS

Subjects	Ave. Err. PS1	Ave. Err. PS2	Gain over PS1		Ave. Err. PS3	Gain over PS1		Gain over PS2		Ave. Err. PS4	Gain over PS1		Gain over PS2		Gain over PS3		Ave. Err. PS4 +PS3	Gain over PS1		Gain over PS2	
			Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain	Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain
1	3.35	2.93	.42	12.5	1.89	1.46	43.6	1.04	35.5	2.40	.95	28.3	.53	1.8	−.51	−27.0	2.15	1.20	35.8	.78	26.6
2	3.61	2.60	1.01	28.0	1.83	1.78	49.3	.77	29.6	1.79	1.82	50.4	.81	31.1	.04	2.2	1.81	1.80	50.0	.79	30.4
3	7.67	2.87	4.80	62.6	2.75	4.92	64.1	.12	4.2	2.19	5.48	71.4	.68	23.7	.56	20.4	2.47	5.20	67.8	.40	13.9
4	5.84	2.84	3.00	51.4	2.04	3.80	65.1	.80	28.1	2.03	3.81	65.2	.81	28.5	.01	0.5	2.03	3.81	65.2	.81	28.5
5	6.33	5.08	1.25	19.7	2.55	3.78	59.7	2.23	50.0	2.37	3.96	62.5	2.71	53.3	.18	7.1	2.46	3.87	61.1	2.62	51.6
6	8.59	3.36	5.23	60.9	1.78	6.81	79.3	1.58	47.0	1.98	6.61	76.9	1.38	41.1	−.20	−11.2	1.88	6.71	78.1	1.48	44.0
7	4.25	3.11	1.14	26.8	2.57	1.68	39.5	.54	17.4	1.86	2.39	56.2	1.25	40.2	.71	27.6	2.22	2.04	48.0	.89	28.6
8	4.17	3.89	.28	6.7	2.17	2.00	48.0	1.72	44.2	1.92	2.25	54.0	1.97	50.7	.25	11.5	2.05	2.13	51.1	1.84	47.3
Ave.	5.48	3.34	2.14	39.1	2.20	3.28	60.0	1.14	34.1	2.07	3.41	62.2	1.27	38.0	.13	5.9	2.13	3.35	61.1	1.21	36.2

ages for the different practice situations is given in Table IV. These data show that the reduction of 2.14 errors on PS2 over PS1 is a reliable statistical difference (5). The reduction of 1.14 errors on PS3 over PS2 is also a statistically reliable difference. The reduction of .13 errors between PS3 and PS4, however, is just slightly better than a chance difference.

TABLE IV

RELIABILITY OF THE DIFFERENCES BETWEEN THE AVERAGES FOR THE DIFFERENT PRACTICE SITUATIONS—EXPERIMENTAL GROUP

	PS1	PS2	PS3	PS4	PS1	PS3+PS4
Average	5.48	3.34	2.20	2.07	5.48	2.13
S.D.	1.82	.76	.35	.22	1.82	.23
S.D. Aver.	.64	.27	.12	.08	.64	.08
Difference	2.14	1.14	.13			3.35
S.D. Diff.	.69	.30	.14			.64
$\frac{D}{\text{S.D. Diff.}}$	3.34	3.80	.09			5.23
Chances in 1,000 of real Diff.	9.99+	999+	538			999+

When the average errors for PS3 and PS4 are combined, the reliability of the difference between this average and that of PS1 is increased.

The data in Tables III and IV show conclusively that there was learning during the 1,000 practice trials without knowledge of results; that there was additional learning during the first 100 trials with knowledge of results; that there was little additional improvement during the second 100 trials with knowledge of results. The failure of these subjects to make substantial gains on PS4 indicates that they had practically reached their maximum performance.

The mass of data in Table III makes it somewhat difficult to compare the performance on PS2, PS3, and PS4 with PS1. The data in Table III, however, are arranged in Fig. 2 so that such comparisons may be made easily.

On examination of Fig. 2, the great individual differences in initial performance and during the practice situations are readily seen. Subjects 1, 2, 4, and 8, who showed the best initial performance, showed the least reduction in errors during

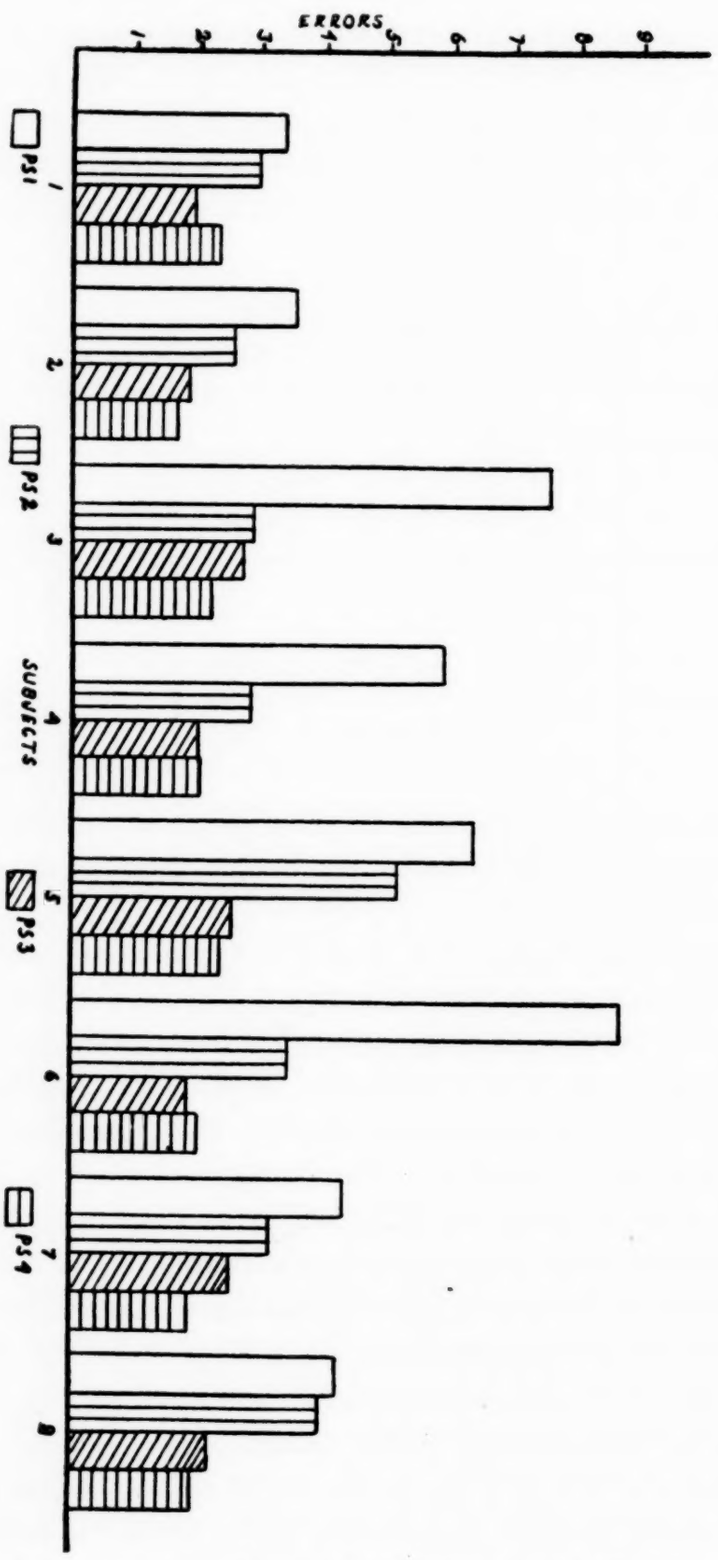

FIG. 2. Average errors for each 50 trials for the 1,300 trials of the experimental group

PS2; while subjects 3, 4, 5, and 6, who registered the largest average error in PS1, showed the greatest reduction in errors during the 1,000 practice trials. It is significant, however, that every subject did materially improve his performance on the 1,000 practice trials. It is also equally significant that each of the eight subjects made additional reduction in errors during the first 100 trials with knowledge of results. The results of PS4, however, do not show improvement for all subjects. Two subjects (1, 6) were less efficient; one subject (4) remained the same; the other five subjects improved slightly.

The performance of the subjects of the experimental group in relation to plus (overlapping) and minus (failure to close) errors is interesting. These data are presented in Table V. Plus errors predominate. This is more than a trend, as it is true for every subject, and the predomination of plus errors is marked. Improvement as far as these subjects are concerned was largely a matter of reducing overlapping. *In computing average errors, signs were disregarded.*

Control Group.—The practice situations for the control group were identical in all respects to those of the experimental group except that they did not have the PS2 trials (the 1,000 practice trials). The average errors made by the subjects of the control group for each of their practice situations is given in Table VI. In addition to the average errors made by each subject for the different practice situations, PS1 performance is compared with that of PS3, PS4, and PS3 + PS4. Gains are shown both in the reduction of errors and the gains in percentages. Reading from top to bottom, the performance of the individual subjects for any practice situation may be examined. Reading from left to right, the reduction in errors and gains in percentages for each practice situation may be studied. The bottom row of figures contains group averages, group reduction in errors, and gains in percentages for the different practice situations. For PS3, individual reduction varied from .68 to 4.79 errors which, expressed in percentages, is a 21.2 to 59.6 percent gain. Individual variations in error and percentage gain for the other practice situation may be determined by examining the data in Table VI.

TABLE V

AVERAGE PLUS AND MINUS ERRORS FOR EACH 50 TRIALS FOR THE EXPERIMENTAL GROUP

Each 50 Trials	Subjects															
	1		2		3		4		5		6		7		8	
	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
1	3.18	.36	3.66	.58	8.20	.22	6.52	.02	7.06	.62	7.96	.40	5.00	.40	4.10	.36
2	1.88	1.28	2.08	.90	6.88	.04	5.00	.14	4.00	.98	8.80	.02	2.76	.34	3.78	.10
3	1.84	1.20	2.00	.40	5.76	.30	2.62	.24	4.62	.60	4.94	.06	3.98	.08	4.14	.10
4	1.84	1.50	1.84	1.16	1.72	.60	1.60	.70	5.28	.68	6.22	.04	3.82	.14	4.10	
5	1.96	.66	1.44	1.56	5.08	.54	3.20	.18	3.70	.74	3.36	.22	3.50	.26	3.70	.06
6	2.94	.44	1.44	1.04	2.36	.80	1.58	.28	3.94	1.08	4.10	.18	3.50	.36	4.34	.08
7	3.90	.12	1.72	.52	2.64	.56	3.54	.42	2.50	1.68	3.22	.24	2.32	.50	4.18	
8	2.26	.60	1.30	.92	.88	1.78	1.94	.52	3.70	1.24	4.08	.02	3.18	.32	3.18	.06
9	2.46	.26	1.42	.34	.60	2.02	1.94	.62	2.84	.70	3.88	.14	3.20	.56	3.48	.10
10	2.08	.60	3.04	.50	1.04	1.18	1.54	.80	4.22	.30	2.82	.28	2.92	.44	4.12	
11	1.72	.48	3.74	.24	.78	1.12	3.04	.28	6.42	.66	2.76	.26	2.66	.18	4.72	
12	1.96	.58	2.60	.34	1.34	1.16	2.58	.40	2.92	.62	2.42	.44	2.14	.52	3.62	.20
13	1.80	.78	2.20	1.40	2.36	.44	2.02	.40	3.44	.88	2.10	.54	2.18	1.10	5.18	.04
14	.78	1.02	1.20	1.74	1.36	.68	2.36	.54	6.70	.14	1.40	.64	1.90	.90	4.06	.08
15	3.52	.46	1.24	1.28	1.36	1.14	2.94	.64	5.10	.40	2.32	.62	2.82	.38	4.18	
16	2.04	.40	.94	1.36	1.56	.88	3.06	.14	3.60	1.38	2.16	.38	2.84	.42	4.92	
17	1.96	.84	.60	1.50	2.50	.54	2.82	.12	3.60	2.64	2.00	.38	1.90	.68	4.34	
18	2.30	.36	.38	1.46	2.68	.36	2.28	.12	4.42	.74	2.06	.38	1.78	1.10	3.18	
19	1.64	1.02	1.58	.78	1.56	1.32	1.46	.58	4.54	.76	2.40	.34	1.24	.72	2.76	.18
20	1.08	1.56	.98	.68	.80	1.48	2.54	.38	2.82	2.00	2.90	.06	1.80	1.50	3.54	.08
21	5.00	.36	1.78	.76	.92	1.38	3.02	.10	2.94	2.32	3.56		1.38	.84	2.44	.16
22	1.16	1.10	2.30	.36	1.00	.82	3.04	.30	2.80	1.94	3.10	.10	1.82	.34	2.48	
23	.60	1.30	1.20	.76	2.34	.88	1.42	.56	1.64	1.18	1.44	.22	2.10	.44	1.86	.32
24	.66	1.22	.80	.90	1.40	.88	1.56	.54	1.10	1.18	1.62	.28	1.74	.86	1.80	.36
25	1.76	1.22	1.00	.68	1.06	1.14	.82	1.34	1.70	.98	1.42	.42	1.26	.76	1.78	.26
26	1.10	.72	1.06	.84	.76	1.42	1.16	.74	.88	1.18	1.78	.34	1.14	.56	1.28	.52
Aver.	2.06	.79	1.68	.88	2.27	.91	2.52	.43	3.71	1.06	3.26	.27	2.50	.57	3.51	.12

TABLE VI

THE AVERAGE ERRORS MADE BY THE SUBJECTS OF THE CONTROL GROUP ON THEIR FIRST 100 TRIALS COMPARED WITH THE 200 TRIALS WITH KNOWLEDGE OF RESULTS

Subjects	Average Errors PS1	Average Errors PS3	Gain over PS1		Average Errors PS4	Gain over PS1		Gain over PS3		Average Errors PS3+PS4	Gain over PS1	
			Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain		Err. Red.	% Gain
9	3.21	2.53	.68	21.2	3.05	.16	5.0	−.52	−20.5	2.79	.42	13.1
10	4.03	2.05	1.98	49.1	1.86	2.17	53.8	.19	9.2	1.95	2.08	51.5
11	7.89	3.90	3.99	50.6	2.90	4.99	63.2	1.00	25.6	3.40	4.49	56.9
12	5.78	3.15	2.63	45.5	2.60	3.18	55.0	.55	17.5	2.87	2.91	50.3
13	6.67	2.41	4.26	62.3	2.07	4.60	69.0	.34	14.1	2.24	4.43	65.7
14	8.03	3.24	4.79	59.6	2.92	4.11	51.2	−.68	−21.0	3.58	4.45	55.4
15	4.10	2.19	1.91	46.6	2.39	1.71	41.7	−.20	−9.1	2.29	1.81	44.2
16	4.41	2.28	2.13	48.3	2.42	1.99	45.1	−.14	−6.1	2.35	2.06	46.7
Ave.	5.52	2.72	2.80	50.7	2.65	2.86	51.8	.07	2.6	2.69	2.83	51.3

The average error reduction for the group in PS3 over PS1 was 2.80, which is 50.7 percent gain. The group reduction for PS4 over PS3 is only .07 error or 2.6 percent. When the average for the 200 trials with knowledge of results is compared with the average for the first 100 trials, 2.83 errors were eliminated. This is an improvement over initial performance of 51.34 percent.

The reliability of the difference between the averages for the different practice situations is given in Table VII. These

TABLE VII

RELIABILITY OF THE DIFFERENCES BETWEEN THE AVERAGES FOR THE DIFFERENT PRACTICE SITUATIONS—CONTROL GROUP

	PS1	PS3	PS4	PS1	PS3+PS4
Average	5.52	2.72	2.65	5.52	2.69
S.D.	1.75	.60	.60	1.75	.54
S.D. Ave.	.62	.21	.21	.62	.19
Difference	2.80	.07		2.83	
S.D. Diff.	.65	.29		.65	
$\frac{D}{\text{S.D. Diff.}}$	4.31	.24		4.35	
Chances in 1,000 of real Diff.	999+	596		999+	

data show that the elimination of 2.80 errors in PS3 over PS1 is a statistical reliable difference (5). The elimination of .07 errors in PS4 over PS3, however, is just slightly better than

a chance difference. When the average error for the 200 trials with knowledge of results is compared with PS1, the reliability of the difference is slightly increased.

The data in Table VI have been organized in Fig. 3 so that both individual and group performance for the different practice situations may be more easily studied and compared. There is great similarity in the shape of the curves in Fig. 2 and Fig. 3. Individual differences in learning are marked;

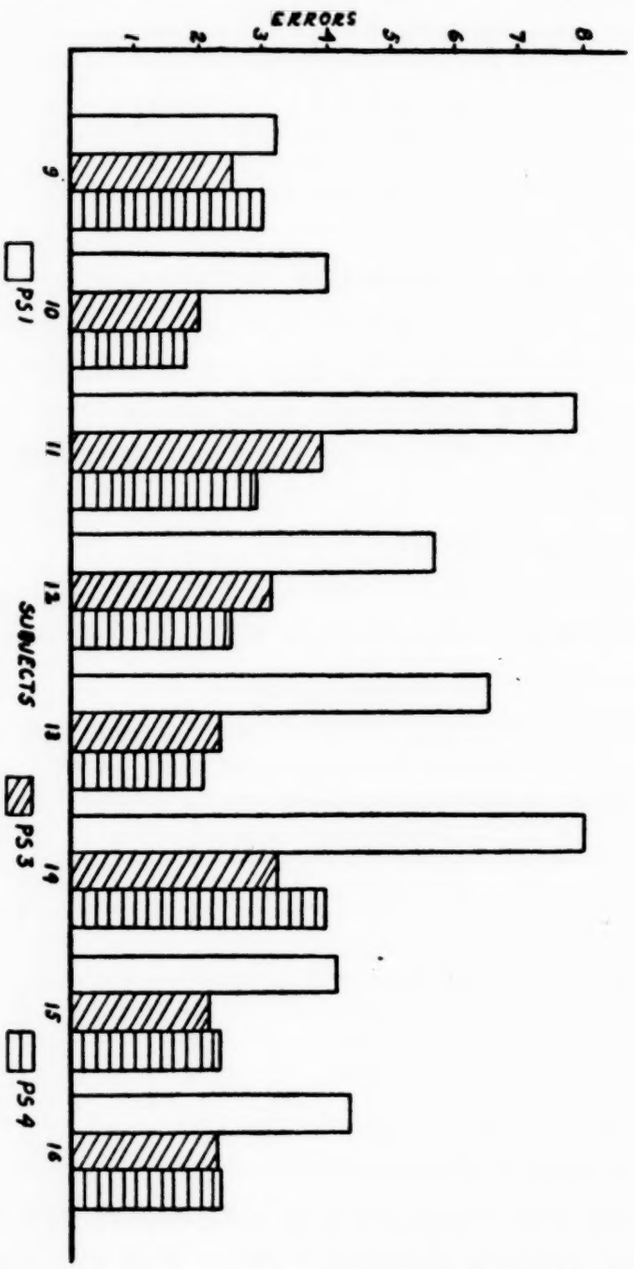

FIG. 3. Average errors for each 50 trials for the 300 trials of the control group.

and we find for the control subjects, as well as for the experimental, the subjects who were most efficient on PS1 (subjects 9, 10, 15, and 16) improved little, while those subjects who showed poorest (subjects 11, 12, 13, and 14) initial performance reduced their errors the most. It is significant, however, that all subjects did improve their performance on PS3 over PS1. A comparison of PS4 with PS3 shows that the second 100 trials with knowledge of results produced little gain over the first 100 trials with knowledge of results. Of the eight subjects, four (subjects 10, 11, 12, 13) improved slightly in PS4 over PS3 and four (subjects 9, 14, 15, and 16) were slightly less efficient. The failure of these subjects to make substantial gains in PS4 would indicate that learning beyond this point through additional practice with knowledge of results would be slight, if present at all.

In studying the plus and minus errors of the subjects of the control group, Table VIII, we find that, as with the members of the experimental group, plus errors predominated. This plus error domination is consistent, being again true for all subjects. Learning, then, for all 16 subjects appears to be largely a matter of reducing overlapping. *In computing average errors, signs were disregarded.*

The data in Tables VI and VII show conclusively that there was learning during the first 100 trials with knowledge of results, but that there was little additional improvement during the second 100 trials with knowledge of results. The failure of these subjects to make substantial gains in PS4 indicates they, like the subjects of the experimental group, had practically reached their maximum performance.

Comparison of Group Learning.—In studying the achievements of the subjects of the experimental and of the control groups, certain important similarities were found. These similarities are as follows:

1. Near identical achievement on the first 100 practice trials.
2. Marked individual differences in achievement.
3. Those who did best on the first 100 trials improved less during subsequent trials.

TABLE VIII

AVERAGE PLUS AND MINUS ERRORS FOR EACH 50 TRIALS FOR THE CONTROL GROUP

Each 50 Trials	Subjects															
	9		10		11		12		13		14		15		16	
	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
1	2.62	.70	4.64	.10	7.76	.54	3.70	2.18	7.42	.82	6.64	.36	4.20	.68	3.78	.32
2	2.24	.86	3.26	.06	7.10	.38	5.44	.24	4.42	.68	9.02	.04	1.80	1.52	4.50	.22
3	1.94	1.02	1.82	.44	3.44	1.14	2.78	.52	1.88	.90	2.66	1.36	1.04	1.16	2.06	.22
4	1.74	.36	1.34	.50	2.58	.64	2.62	.38	1.32	.72	1.42	1.04	.66	1.52	1.44	.84
5	3.18	.28	1.56	.46	2.12	.70	2.10	.38	1.30	.82	2.36	1.64	1.02	1.46	2.82	.34
6	1.86	.78	1.02	.68	2.16	.82	2.12	.60	1.26	.76	2.16	1.68	.58	1.72	1.38	.30
Ave.	2.26	.67	2.28	.38	4.19	.70	3.13	.72	2.93	.78	4.04	1.02	1.55	1.34	2.66	.37

4. Those who performed poorest during the first 100 trials improved most during subsequent trials.
5. Definite increase in efficiency during the first 100 trials with knowledge of results.
6. Slight group improvement during the second 100 trials with knowledge of results over PS3.
7. More plus errors (overlapping) than minus errors (failure to close).
8. That improvement in performance in this type of acquisition is largely a matter of reducing plus errors.

In making a further comparison of learning efficiency between the two groups, only the group averages for each practice situation are considered. These data are given in Table IX. Although the groups were practically identical in performance following 100 attempts to draw a circle without knowledge of results, after the experimental group were given 1,000 additional trials without knowledge of results, the two groups differed by 2.18 errors. This difference is a statistically reliable difference (Table IV). One of the problems in this investigation was to measure the learning (objective) resulting from the practice of drawing circles without knowledge of results. This has been done, as 1,000 practice trials drawing circles without knowledge of results produced a 39.1 percent gain in efficiency.

The second problem was to investigate the latent learning resulting from practice in drawing circles without knowledge of results. If, through practice in drawing circles without knowledge of results, skills are accumulated which do not show objectively until there is knowledge of results, these skills should show when the group averages for the first 100 trials with knowledge of results are compared.

Comparing the performance of the two groups with their initial performance (Table IX) shows that the experimental group gained 60 percent and the control group 50.7. The experimental group was now only 9.3 percent better in performance than the control group. The experimental group had gained 39.1 percent during the 1,000 trials without knowl-

TABLE IX

A Comparison of the Average Errors Made by the Experimental Group on the Different Practice Situations with the Average Errors Made by the Control Group

	Average Errors PS1	Average Errors PS2	Gain over PS1		Average Errors PS3	Gain over PS1		Gain over PS2		Average Errors PS4	Gain over PS1		Gain over PS2		Gain over PS3		Average Errors PS3 +PS4	Gain over PS1		Gain over PS2	
			Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain	Err. Red.	% Gain		Err. Red.	% Gain	Err. Red.	% Gain
Exp. . . .	5.48	3.34	2.14	39.1	2.20	3.28	60.0	1.14	34.1	2.07	3.41	62.2	1.27	38.0	.13	5.9	2.13	3.35	61.1	1.21	36.2
Con. . . .	5.52				2.72	2.80	50.7			2.65	2.86	51.8			.07	2.6	2.69	2.83	51.3		
Diff. . . .	.04				.52	.48	9.3			.58	.55	10.4			.06	3.3	.56	.52	9.8		

edge of results; so, instead of a sudden spurt in improvement resulting from skills accumulated during practice without knowledge of results, these trials with knowledge of results produced less improvement than was achieved during the PS2 trials.

This slowing down in the learning process would indicate the absence of latent learning. A further indication of the lack of latent learning is shown when the improvement of the control group in PS3 is compared with that of the experimental. At the beginning of the PS3 trials, the control group's average error was 2.18 more than that of the experimental, but after the first 100 trials with knowledge of results, the control group's average error was only .52 more than that of the experimental group. The difference of .52 error, however, is very close to a statistically reliable difference (Table X). In

TABLE X

THE RELIABILITY OF THE DIFFERENCES BETWEEN THE AVERAGES OF THE EXPERIMENTAL AND CONTROL GROUP FOR THE DIFFERENT PRACTICE SITUATIONS

	PS1		PS3		PS4		PS3+PS4	
	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control
Average	5.48	5.52	2.20	2.72	2.07	2.65	2.13	2.69
S.D.	1.82	1.75	.35	.60	.22	.60	.23	.54
S.D. Ave.	.64	.62	.12	.21	.08	.21	.08	.19
Difference	.04		.52		.58		.56	
S.D. Diff.	.89		.24		.22		.21	
$\frac{D}{\text{S.D. Diff.}}$	.05		2.17		2.64		2.66	
Chances in 1,000 of real Diff.	520		987		996		996	

100 trials with knowledge of results, the control group improved almost as much as the experimental group had during 1,100 trials, 1,000 of which were practice without knowledge of results. Instead of the difference between the average performance of the two groups becoming greater at the end of the PS3 trials, it was much reduced.

If skills in drawing circles had accumulated during the

1,000 practice trials without knowledge of results, but had not shown already objectively, these skills did not show in a sudden and marked increase in efficiency during the first 100 trials with knowledge of results. The net gain in efficiency in circle drawing in the 1,000 trials without knowledge of results was approximately only nine percent, not a very high reward when the time and effort expended on the 1,000 trials is considered. No latent learning is discovered in comparing the PS3 averages; but this comparison does show that, although there is learning through practice without knowledge of results, approximately the same degree of efficiency may be obtained with much less effort through practice with a knowledge of results. In 1,000 practice trials without knowledge of results, the experimental group gained 39.1 over their initial efficiency; the control group on 100 trials with knowledge of results gained 50.7 percent.

Perhaps 100 trials drawing circles with knowledge of results is not sufficient practice to bring out the skills accumulated by the experimental group during the 1,000 practice trials without knowledge of results. A comparison of the two groups on the PS4 trials will show whether this is true. This comparison (Table IX) shows, however, that the difference between the averages of the two groups remained practically the same. Both groups gained slightly in efficiency during the PS4 trials over the PS3 trials. *Thus, there is no indication of latent learning.* The failure to make substantial gain during the PS4 trials indicates that both groups have practically reached their maximum efficiency and additional trials in drawing circles with knowledge of results would probably not bring about an increase in efficiency.

The failure of the control group to attain the same degree of efficiency as that of the experimental group (Table IX) during the 200 trials with knowledge of results, and the high reliability of the difference between the averages of the two groups (Table X) at the end of the PS3 and PS4 trials indicates that the approximate 10 percent gain in efficiency of the experimental group cannot be equaled by the control group provided both groups are given the same amount of additional practice with knowledge of results.

An easier comparison of the learning of the two groups can be made by examining Fig. 4. The solid bar represents the average errors made by the experimental group in the different practice situations, and the open bar, the control. This graph

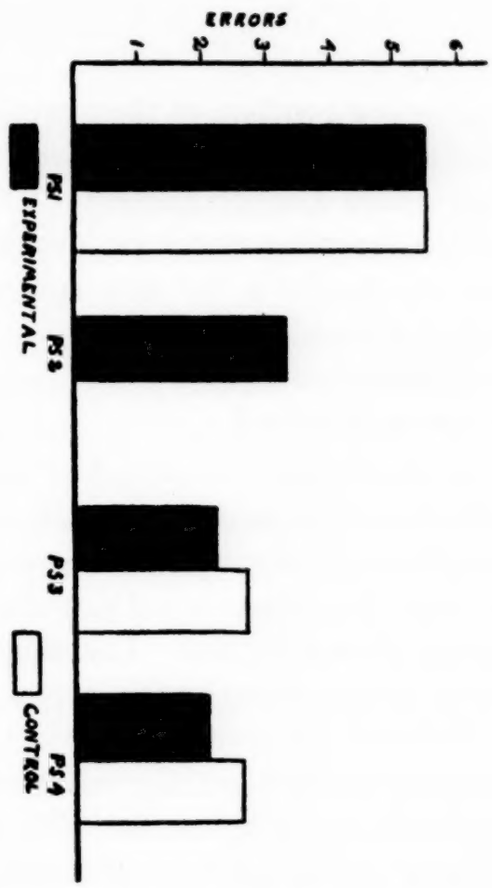

FIG. 4. A comparison of the average gains made by the experimental group in the different practice situations with the average gains made by the control group.

shows vividly that if two groups of equal ability are practiced in drawing circles on the rotameter, one group given 1,200 trials (1,000 trials without knowledge of results, the last 200 with knowledge of results) and the other group only 200 trials (all these trials with knowledge of results), the improvement due to the 1,000 practice trials without knowledge of results produced, approximately, only a 10 percent gain in efficiency over the group that had only the 200 trials with knowledge of results.

Interpretation of Results

The degree of efficiency achieved through practice without knowledge of results seems to be affected by a number of major factors. Therefore, before the results of this study can be adequately understood, the following factors need careful consideration:

1. Whether the learning consists of the acquisition of entirely new responses or whether it consists of the improvement of responses already present.
2. The nature of the learning-motor or ideational.
3. The technique of acquiring the new responses or the improvement of the old ones.
4. The fineness of measurements.
5. The sense or senses involved.

1. Learning, in this study, consisted of the improvement of responses which the subjects could already make. Each of the 16 subjects could rotate the stylus of the rotameter to draw a circle the first trial, but they could not draw accurately a circle with a radius of two inches. Learning was measured by the reduction in errors resulting from practice in rotating the stylus.

2. The decrease in error resulting from practice without knowledge of results in drawing circles with a radius of two inches was apparently the acquisition of motor skill. Verbalization, mental images, and other forms of ideational behavior may have been involved, but such data were not secured. The objective data show only the improvement of motor skills.

3. The technique of improving the ability to draw circles with a radius of two inches without knowledge of results consisted of having each subject of the experimental group make 1,000 attempts to stop the stylus of the rotameter at exactly the same point from which he started. There were no external cues nor standards that would help the subject verify the accuracy of his response. The practice with knowledge of results was for the purpose of: (*a*) investigating latent learning, and (*b*) comparing the extent of learning through practice without and with knowledge of results.

4. The use of a mechanical device, the rotameter, to control the practice trials made accurate objective measurements possible.

5. Two senses were directly involved, tactual and kinaesthetic. Seeing, hearing, tasting, and smelling were eliminated. The sensations arising from pressure on the fingers grasping the stylus during the movement of the needle from one position to another would provide a method for verification of the accuracy of the response. Some degree of tactual acuity was undoubtedly present at the beginning of the 1,000 practice trials without knowledge of results. Improvement in performance, then, could have been in part either an increase in sensory acuity or an increase in attention given to the tactual sensations. It can also be assumed that some degree of kinaesthetic acuity was present at the beginning of the practice trials. Whether the reduction in errors was in part a development of new kinaesthetic responses or merely the giving of more attention to those sensations already present is an interesting question.

The results of this study do not permit an analysis of improvement into units of tactual and kinaesthetic development. The data strongly suggest, however, that neither new tactual nor new kinaesthetic responses were acquired; but, rather, a more efficient use of those which were already present. Considerable evidence points to this interpretation: (1) Improvement consisted largely of the reduction of plus errors (overlapping). This would indicate that, as practice progressed, more attention was given to tactual and kinaesthetic responses. (2) Marked individual differences in initial efficiency and in final achievement may be interpreted as differences in the quality and quantity of the tactual and kinaesthetic responses and the degree of attention given to these responses. (3) The failure to make substantial gains on the second 100 trials with knowledge of results would indicate that efficiency in drawing circles under the experimental conditions of this study would not go beyond that of the tactual and kinaesthetic responses already present at the beginning of the experiment.

Summary

To investigate the latent learning resulting from practice in drawing circles without knowledge of results, and to measure the objective learning resulting from such practice, two groups of eight subjects, each of equal ability, were tested on the rotameter. One group, the experimental, was given 1,000 practice trials without knowledge of results and 200 trials with knowledge of results. The other group, the control, was given only the 200 practice trials with knowledge of results. The performance of each group in the different practice situations was studied, and the efficiency of the experimental group for each practice situation was compared with that of the control group. Some of the more important results of the study are the following:

1. No measurable amount of latent learning was accumulated during the 1,000 trials without knowledge of results.
2. Efficiency was increased 39.1 percent during 1,000 practice trials in drawing circles on the rotameter without knowledge of results.
3. A few practice trials in drawing circles with knowledge of results produce greater efficiency in performance than do many trials without knowledge of results.
4. Increased efficiency developed through practice without knowledge of results is in part a permanent gain that is not overcome by the group that did not have the 1,000 practice trials without knowledge of results.
5. There were marked individual differences in initial performance and in amount of gain in the different practice situations.
6. Subjects with the best initial performance improved less on subsequent trials.
7. Subjects with the poorest initial performance improved most on subsequent trials.
8. A maximum degree of efficiency had practically been reached by the end of 200 trials with knowledge of results.
9. Plus errors dominated throughout the trial series.

10. Learning was largely a matter of reducing overlapping (plus errors).

(Manuscript received August 17, 1935)

References

1. Arps, G. F., Work with knowledge of results versus work without knowledge of results. *Psy. Monogs.*, 1920, **28**, 1–41.
2. Blodgett, H. C., The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Pub. in Psychol.*, 1929, **4**, 113–34.
3. Book, W. F. and Norvell, L., The Will to Learn. *Ped. Sem.*, 1922, **29**, 305–62.
4. Crafts, L. W. and Gilbert, R. W., The effect of knowledge of results on maze learning and retention. *J. Educ. Psychol.*, 1935, **26**, 177–87.
5. Garrett, H. E., *Statistics in Psychology and Education.* N. Y.: Longmans, Green, 1926.
6. Hall, C., A repetition of Thorndike's experiment on improvement in the estimation of lengths. *J. Gen. Psychol.*, 1933, **9**, 238–41.
7. Judd, C. H., Practice without knowledge of results. *Psychol. Rev. Monog. Supp.* 1905, **7**, No. 29, 185–94.
8. Ross, C. C., An experiment in motivation. *J. Educ. Psychol.*, 1927, **18**, 337–346.
9. Smith, F. D., Repetition without knowledge of results as a factor in learning. *Psychol. Bull.*, 1933, **30**, 673.
10. Spencer, L. T., Effects of practice without knowledge of results. *Amer. J. Psychol.*, 1923, **34**, 107–111.
11. Thorndike, E. L., *Human Learning.* N. Y.: Century, 1931.
12. Tolman, E. C., and Homzik, C. H., Introduction and removal of reward, and maze performance in rats. *Univ. Calif. Pub. in Psychol.*, 1930, **4**, 257–275.

CERTAIN SOMATIC ACTIVITIES IN RELATION TO SUCCESSFUL AND UNSUCCESSFUL PROBLEM SOLVING[1]

BY MYRON S. CLITES

Indiana University

PART I

Inter-Wink Intervals During Problem Solving

At present there is some diversity of opinion as to the relationship between mental work and the extraneous activities which are sometimes observed. If it should be found that there is less muscular tension, glandular activity and incipient movement during problem solving, this would furnish evidence for supposing that when one type of work is in progress the other type of work is diminished. But, if it should be found that there is increase in muscular tensions and in reflex activities during problem solving, then it might be inferred that (1) somatic activity accompanies problem solving, or (2) problem solving consists in whole or in part of somatic activity. Before any of these theoretical questions can be answered, some basic experimental work on the topic needs to be done as the evidence at this time is somewhat conflicting. Woodworth [**22**] suggests that the involuntary muscular tensions of jaw, brow, neck, hands, body, and legs which so regularly accompany mental effort, may have a genuine facilitating effect upon mental activity. But he adds, it has been found that adding and memory tasks are better performed in a reclining than in a standing position.

The first question to be attacked is to determine whether or not there is an increase in various types of bodily performance during problem solving as compared with periods of mental rest. Then it should be determined whether or not

[1] Written under the direction of R. C. Davis. Subsequent articles will deal with other somatic activities.

there is any correspondence in such physiological factors as the amount of reflex activities and the muscular tensions being exerted and the success or failure of the mental work being done.

If it should be found that successful mental work is accompanied by a greater decrease in bodily performance than is unsuccessful work, it might be concluded that effectiveness of mental work consists of a redirecting of extraneous activity while mental activity is in progress. But, if it should be found that successful mental work is accompanied by an increase in bodily performance, the conclusion might be reached that mental effectiveness consists of becoming widely active when a mental task is presented.

However, before we can explain the relationship of success in mental work to bodily tensions and reflexes, it is necessary to determine (1) whether people who solve problems successfully have increased or decreased muscular tensions and reflexes during mental work, or (2) whether people who persevere but fail to solve problems have an increase or decrease in these phenomena, and (3) what happens with regards to these activities when people 'give up.'

Specifically, this research is concerned with the relation between problem solving and a function which is a phasic activity and can be controlled in part at the discretion of the subject, *i.e.*, the rate of involuntary wink.

Ponder and Kennedy [**20**] conclude from their investigation that the eye wink is under the control of the central nervous system and that it is accelerated by such factors as excitement, 'anger,' and cross-examination in court. They hypothecate that the eye wink is a sort of relief mechanism whereby nervous energy, when otherwise unutilized as in emotional conditions, passes into a highly facilitated path. On the basis of observations made on street cars and in a library, Ponder and Kennedy also conclude that in the case of intense mental work the energy is consumed and there is no need of drainage by way of a facilitated path. This hypothesis is worthy of further investigation.

In the field of muscular tensions and muscular movements,

W. F. Book [2] found that when a subject is striving hard to overcome tendencies to error in typing he uses extra force in hitting the keys and space bar. Morgan's [19] finding is comparable for he says that when a subject is annoyed during work that this extra stimulation calls for greater expenditure of energy and more physical work is actually done. These workers report what happens in the case of activities which are indispensible to the work in hand. It would be interesting to determine what happens in the case of activities which seemingly are not so closely related to the task in progress. Thorson [21] concludes that movements of the tongue are not universal in internal speech or verbal thought. This leads to the further question as to whether or not there might be other movements which are related to successful and unsuccessful problem solving.

Clark [3] says that an individual is noticeably quiet when he is actually thinking, for a pause in movement is usually apparent just before the thinker indicates he has reached a novel conclusion. Possibly this pause in movement is due to an increase or decrease in muscular tensions.

Freeman [6] concludes that tonus decreases with practise in learning nonsense syllables. He says that cortical centers predominantly involved in mental activity have relatively high excitation thresholds and by appropriate lowering of these thresholds, they can be made to respond to otherwise inadequate stimuli. Such threshold-alteration is accompanied by irradiation from lower centers, particularly those involved in muscular contraction. If learning nonsense syllables and problem solving require the same type of activity it would follow that during problem solving there would be an increase in muscular contraction. Ford [5] says that muscle tonus is an index of the degree of attention because he finds that writing-pressure and reaction-time show a reliable positive correlation of .55. If attention is necessary for successful problem solving one should expect that it would be accompanied by a higher muscular tonus than is unsuccessful work.

Jones [16] found by questionnaire that eminent men usu-

ally assumed a reclining or half reclining position when doing their most effective mental work. He assumed that the reclining position was accompanied by muscular relaxation but this is in need of experimental evidence.

Luria [**18**] reports that when a person guilty of crime is required to squeeze a pneumatic bulb while simultaneously responding with free association words there is a harder and more irregular squeeze than in the case of innocent individuals. The hard and irregular squeeze was elicited in the case of those about to take an important examination. Such findings would lead one to suppose that failure in a problem solving situation would likewise be accompanied by some sort of disturbance in somatic activity.

Freeman and Katzoff [**7**] find that the correlation between the score on Cason's Common Annoyance Test and tension as expressed by pressure on a pneumatic bulb while being disturbed during mental work is positive .49. Duffy [**4**] finds that dynamograph records of pressure correlate −.47 with tapping performance and −.48 with discriminative performance in the case of children ranging from two years eleven months to three years and ten months of age. The question arises as to what would happen with regards to the muscular pressure during such mental activity as successful and unsuccessful problem solving.

Lombard [**17**] reports that the amplitude of the knee jerk is decreased during physical and mental fatigue. He finds that it can be reënforced by irritation of the skin, simultaneous voluntary actions, stimuli exciting attention, rapid mental multiplication, reading a stirring poem, and interesting or lively music. Since Lombard found differences in the amplitude of the knee jerk during different kinds of activity, it seems that probably there might be a difference in various somatic activities when people succeed or fail in problem solving situations.

In a series of studies dealing with electrical measurements of neuromuscular activity during mental activity, Jacobson [**8–14**] finds that such a mental act as imagination of performance by some part of the body is accompanied by tensions in

that part. In our investigation an attempt is made to determine the relationship of another type of mental performance, namely, problem solving, to muscular movements and tensions.

Bills [1] reports that squeezing a dynamometer is accompanied by increased efficiency in learning nonsense syllables, learning paired associates, in adding and in calling names of letters on exposed cards. This is the converse of the problem solving experiment we are conducting which proposes to compare somatic activities normally arising in the course of good and poor problem solving.

In an attempt to determine whether or not the success or failure of mental work is in any way related to the frequency of eye wink, we are devoting this section of this research to an investigation of the length of inter-wink intervals before problems solving, during persistent but unsuccessful work, during successful solution, during work followed by giving up, during reporting of solution, and during rest after mental work.

Experimental Procedure

The purpose of this section is to determine (1) the relationship of rate of winking to successful and unsuccessful mental work as compared with periods of rest for the respective groups, and (2) whether or not there is any difference in rate of winking during mental work as compared with mental rest in the case of those who persevere but fail and those who give up.

The experimenter asked the subject to be seated and to refrain from thinking of anything specifically. A long record kymograph fed from a bolt of wrapping paper and travelling at the rate of .3 cm per second was set going. The occurrence of winks was recorded by E's tapping a telegraph key while the time was taken from a pendulum timer swinging at one-second intervals. Various mechanical devices for recording the winks were tried but the subjects too often suspected what was being done and the subjective factor vitiated the results. E. practised recording the occurrence of winks until he could do it accurately. Another key was used to indicate

the beginning and end of the periods of 'rest,' problem reading, problem solving, solution reporting, and 'rest' after the problem. The rest period after one problem served as the rest period before the next problem. Length of inter-wink intervals to the nearest second was used as the basis of computation. The problems used were the three water dipping problems found at the superior adult level of the Binet Test.

In case a subject failed to get the correct solution, the experimenter stopped the kymograph and carefully explained to the subject how to solve the problem. Then E. told S. not to bother thinking of the problem any more as the purpose was merely to detect differences occurring during mental work and mental rest. After approximately a minute the record-taking was resumed.

TABLE I

NUMBER OF SECONDS BETWEEN WINKS DURING THE PERIODS SPECIFIED FOR THOSE WHO SUCCEEDED

	Rest I	E. Reading I	Subject Solving I	Subject Reporting I	Rest II	Rest II	E. Reading II	Subject Solving II	Subject Reporting II	Rest III	Rest III	E. Reading III	Subject Solving III	Subject Reporting III	Rest IV
Number subjects	33	33	33	33	33	23	23	23	23	23	31	31	31	31	31
Ave. I.-W. interval	8.4	3.1	3.8	3.0	7.7	7.0	2.7	3.2	3.6	7.4	6.2	3.0	4.1	3.9	7.9
Diff.															
S.D.(D.)															
With Rest I		5.2	4.1	5.2											
With Rest II		5.0	4.1	4.6			3.8	3.3	3.0						
With Rest III							3.4	3.5	2.8			4.4	3.0	3.6	
With Rest IV												3.4	2.7	3.6	

TABLE II

NUMBER OF SECONDS BETWEEN WINKS DURING THE PERIODS SPECIFIED FOR THOSE WHO FAILED

	Rest I	E. Reading I	Subject Solving I	Subject Reporting I	Rest II	Rest II	E. Reading II	Subject Solving II	Subject Reporting II	Rest III	Rest III	E. Reading III	Subject Solving III	Subject Reporting III	Rest IV
Number subjects	14	14	14	14	14	25*	25	25	25	25	21	21	21	21	21
Ave. I.-W. interval	5.7	4.4	3.1	—	4.7	7.8	3.2	3.1	—	5.0	4.9	3.5	3.2	—	4.7
With Rest I		1.35	3.4												
With Rest II		.3	2.8				5.1	4.9							
With Rest III							2.6	3.6				1.6	2.6		
With Rest IV												1.6	2.1		

* Eleven of these subjects had succeeded on previous problem; note the longer inter-wink intervals. In the other cases, people who succeeded had longer inter-wink intervals during rest than did those who failed.

For the subjects who succeeded in solving these Binet problems the inter-wink intervals had a very definite tendency to be shorter during the experimenter's reciting of the problem, during the subject's solving time and during the reporting of the solution than they were during the rest periods before and after the problems. The reliability quotients ($D./S.D._{D.}$) range from 2.7 to 5.2. Thus, people who succeed in solving the problems wink more rapidly while they are being presented the problem, while they are solving it and while they are reporting their solution than they do during the periods of no mental work.

Failing subjects show a definite tendency to wink more rapidly during the time they are working on the problem than they do during rest. The reliability quotients range from 2.1 to 4.9. However, in the case of failing subjects the inter-wink intervals during the reading of the problems are not statistically shorter than they are during rest except in the case of the second rest versus the reading of the second problem. But, it will be noted that 14 of these people had succeeded in solving the first problem and 11 had not. If only the 11 who failed to get both problems are considered, the inter-wink intervals are not statistically shorter during the reading of the second problem than they are during the rest period preceding this problem.

The group which failed was divided into two sections, those who gave up and those who persevered to the end of the solving period. Perseverence was said to prevail when the subject made comments during solving which indicated he was solving, when he attempted a report or when the experimenter stopped the subject. 'Giving up' was indicated by the subject's informing the experimenter that he could not solve the problem. All the problems for those who gave up and for those who persevered, respectively, were thrown together. (See Table III.)

In the case of those who persevered, the inter-wink intervals during the reading of the problem by the experimenter are not definitely shorter than during the preceding or succeeding rest periods. Also, the inter-wink intervals

TABLE III

THE LENGTH OF INTER-WINK INTERVALS (TO THE NEAREST .1 SECOND) OCCURRING DURING REST, DURING READING OF PROBLEM, DURING SOLVING, AND DURING REPORTING FOR THOSE WHO 'GAVE UP' AND FOR THOSE WHO PERSEVERED. ALL THE PROBLEMS FOR 'GIVING-UP' SUBJECTS AND FOR THOSE WHO PERSEVERED, RESPECTIVELY, WERE THROWN TOGETHER

(*A*) *Those Who 'Gave Up'*

	First Rest	E. Reads Problem	S. Attempts Solution	Attempts Report	Final Rest
No. subjects	36	36	36	—	36
Ave. length of I.-W. intervals	6.5	3.7	2.7	—	5.0
Difference					
S.D. (Diff.)					
With first rest		4.0	4.3		
With final rest		2.4	5.4		

(*B*) *Those Who Persevered*

	First Rest	E. Reads Problem	S. Attempts Solution	Attempts Report	Final Rest
No. subjects	23	23	23	—	23
Ave. length of I.-W. intervals	4.7	4.1	3.5	—	4.4
Difference					
S.D. (Diff.)					
With first rest		.57	2.3		
With final rest		.3	1.8		

which averaged 3.5 seconds in length during the attempted solution have only a definite tendency (instead of a certain tendency) to be shorter than those inter-wink intervals occurring during the preceding and succeeding rest periods which averaged 4.7 seconds and 4.4 seconds, respectively.

In the case of the subjects who gave up the inter-wink intervals occurring during reading and during solving are definitely shorter than those wink intervals occurring during the preceding and succeeding rests.

Differences in Winking During Different Parts of Solving Period

Since no way could be found to have the subject indicate when he was succeeding, we adopted the plan of computing the length of inter-wink intervals during the first two-thirds and during the last one-third of the solving period. (See Table IV.) The same computations were made for the first

TABLE IV

COMPUTATIONS FOR THE FIRST TWO-THIRDS VS. THE LAST ONE-THIRD OF THE SOLVING PERIOD

	No. Subject	Mean I.-W. Intervals	S.D. (dist.)	S.D. (mean)	r	S.D.$_r$	Diff. / S.D.D.
Subjects succeeded:							
First .67 of solving	89	4.00	3.6	.40	+.75	±.046	1.0
Last .33 of solving	89	4.35	4.6	.50			
Subjects gave up:							
First .67 of solving	36	2.90	2.4	.40	+.93	±.014	3.3
Last .33 of solving	36	3.40	2.5	.40			
Subjects persevered:							
First .67 of solving	23	3.90	4.2	.88	+.83	±.033	1.2
Last .33 of solving	23	3.30	2.6	.54			

.8 and the last .2 of the solving period. (See Table V.) The correlation [2] of the length of inter-wink intervals for successes during the first two-thirds of the solving period and the last one-third of the solving period is positive .75. For those who gave up the corresponding correlation was positive .93; and for those who persevered, the correlation was positive .83.

TABLE V

COMPUTATIONS FOR THE FIRST .8 VS. THE LAST .2 OF THE SOLVING PERIOD

	No. Subject	Mean I.-W. Intervals	S.D. (dist.)	S.D. (mean)	r	S.D.$_r$	Diff. / S.D.D.
Subjects succeeded:							
First .8 of solving	89	3.7	3.0	.32	+.61	±.067	2.50
Last .2 of solving	89	4.7	4.7	.50			
Subjects gave up:							
First .8 of solving	36	2.9	2.1	.35	+.86	±.276	1.50
Last .2 of solving	36	3.3	3.0	.50			
Subjects persevered:							
First .8 of solving	23	3.9	3.9	.80	+.61	±.067	.48
Last .2 of solving	23	3.6	3.5	.73			

These high correlations indicate that any group tendencies are likewise characteristic of the individuals composing the group. The reliability quotients representing the differences between the first two-thirds and the last one-third of the solving period indicate that there is a reliable difference in the

[2] Rank Difference method of computing correlation was used.

case of those who gave up, but not in the case of those who persevered but failed and those who succeeded.

The correlations for the length of inter-wink intervals during the first .8 of the solving period vs. the last .2 of this period are +.61 for successful subjects, +.86 for those who gave up and +.61 for those who persevered. These high correlations indicate that any group tendencies are likewise characteristic of the members composing the groups. The reliability quotients representing the differences in lengths of inter-wink periods between the first .8 and the last .2 of the solving period are 2.5 for successes, 1.5 for those who gave up and .48 for those who persevered but failed.

The computations for the first two-thirds vs. the last one-third of the solving period and for the first .8 and the last .2 indicate that: (1) successful subjects tend to decrease their rate of winking as they come to the end of the solving period, (2) the winking rate for those who give up tends to decrease during the third quarter of the solving period, (3) the winking rate for those who persevere but fail is practically constant during time of attempted solution. It seems, then, that the longer inter-wink intervals or slower winking rate is characteristic of successful mental work.

Significant Differences Between Successful and Failing Subjects Are During the Rest Periods

During the periods of problem presentation, problem solving, and reporting of solution, the absolute differences between the length of inter-wink intervals for successful subjects and for failing subjects are small and not reliable. But, during the periods of rest, the inter-wink intervals tend to be longer for successful than for failing subjects. (See Tables VI and VII.)

In all cases except one there is a definite tendency for the inter-wink intervals to be longer for successful subjects during rest than they are for failing subjects. The reliability quotients range from 1.3 to 2.0 except in one case in which the quotient is .47 in the opposite direction. This one case is during the rest period before the second problem and it will

TABLE VI

DIFFERENCES BETWEEN LENGTH OF INTER-WINK INTERVALS DURING THE REST PERIODS BEFORE AND AFTER EACH PROBLEM FOR SUCCESSFUL AND FAILING SUBJECTS

	Length of I.-W. Intervals	Diff.	S.D. (means)	Diff. / S.D.D.	Chances per Thousand of Difference
Rest I:					
Success	8.4	2.7	1.1	2.0	977
Failure	5.7	—	.8	—	—
Rest II before Prob. II:					
Success	7.0	−.8	1.2	—	—
Failure	7.8	—	1.2	.47*	681
Rest III before Prob. III:					
Success	6.2	1.3	.5	1.3	904
Failure	4.9	—	.9	—	—
Rest II after Prob. I:					
Success	7.7	4.0	1.1	3.2	999
Failure	3.7	—	.6	—	—
Rest III after Prob. II:					
Success	7.4	2.4	1.3	1.7	956
Failure	5.0	—	.7	—	—
Rest IV after Prob. III:					
Success	7.9	3.2	1.6	1.8	964
Failure	4.7	—	.8	—	—

* Slight reversal.

Computation by throwing all initial rests together: $\frac{.023 \times .096}{.319} = .007$ vs. real difference, or 993 chances per thousand that the difference is real.

Computation by throwing all final rests together: $.001 \times .044 \times .36 = .000001584$ vs. real difference, or 999,998 chances in 1,000,000 that the difference is real.

be seen by referring to Table II that there are 14 subjects who failed to solve the first problem and 25 who failed to solve the second. If the inter-wink intervals for only the subjects who failed to solve both problems are considered, the difference at this point is significant and in the same direction as in the other cases of rest. When the chances per thousand against a real difference are multiplied, we find that there is a very definite tendency for the inter-wink intervals to be longer during all the rest periods for successful than for failing subjects.

When the inter-wink intervals during the rest periods before each of the problems are considered in the case of successful subjects and those who gave up, we find a tendency

TABLE VII

DIFFERENCES BETWEEN LENGTH OF INTER-WINK INTERVALS DURING THE REST PERIODS BEFORE AND AFTER EACH OF THREE PROBLEMS FOR SUCCESSFUL SUBJECTS AND SUBJECTS WHO GAVE UP

	Length of I.-W. Intervals	Diff.	S.D. (means)	$\frac{\text{Diff.}}{\text{S.D.D.}}$	Chances per Thousand of Real Difference
Rest I before Prob. I:					
Success	8.4	1.9	1.10	1.3	903
Gave up	6.5	—	.93	—	—
Rest II before Prob. II:					
Success	7.0	.5	1.20	.3	613
Gave up	6.5	—	.93	—	—
Rest III before Prob. III:					
Success	6.2	−.3	.50	—	—
Gave up	6.5	—	.93	.3*	618
Rest II after Prob. I:					
Success	7.7	2.7	1.10	2.2	986
Gave up	5.0	—	.53	—	—
Rest III after Prob. II:					
Success	7.4	2.4	1.30	1.7	955
Gave up	5.0	—	.53	—	—
Rest IV after Prob. III:					
Success	7.9	2.9	1.60	1.7	955
Gave up	5.0	—	.53	—	—

*Slight reversal.

Computations by throwing all initial rests together: $\frac{.097 \times .618}{.618} = .097$ chance vs. real difference or 903 chances per 1000 that a real difference exists.

Computation by throwing all final rests together: $.014 \times .045 \times .045 = .000028350$ chance vs. a real difference, or 999,972 chances in 1,000,000 of a real difference.

for the successful subjects to have longer inter-wink intervals but this tendency is not so definite as it is in the rest periods. It seems, then, that when a problem is being presented to subjects who succeed they tend to wink less during the presentation than do subjects who do not persevere in the working of the problem. But when the rest periods after the problem situation are considered, we find that the chances are very remote that the successful subjects wink more rapidly than do those who gave up.

From the foregoing tables it can be seen that those who succeed in solving the problems wink more slowly during rest and then they materially increase their winking rate during mental work. The winking rate for successful and

TABLE VIII

DIFFERENCES BETWEEN LENGTHS OF INTER-WINK INTERVALS DURING THE REST PERIODS BEFORE AND AFTER EACH OF THE PROBLEMS FOR SUCCESSFUL SUBJECTS AND SUBJECTS WHO PERSEVERE BUT FAIL

	Length of I.-W. Intervals	Diff.	S.D. (means)	$\frac{\text{Diff.}}{\text{S.D.D.}}$	Chances per Thousand of Real Difference
Rest I before Prob. I:					
Successes	8.4	3.7	1.10	3.1	999
Persevered	4.7	—	.43	—	—
Rest II before Prob. II:					
Successes	7.0	2.3	1.20	2.0	977
Persevered	4.7	—	.43	—	—
Rest III before Prob. III:					
Successes	6.2	1.5	.50	2.3	990
Persevered	4.7	—	.43	—	—
Rest II after Prob. I:					
Successes	7.7	3.3	1.10	2.3	990
Persevered	4.4	—	.50	—	—
Rest III after Prob. II:					
Successes	7.4	3.0	1.30	2.2	986
Persevered	4.4	—	.50	—	—
Rest IV after Prob. III:					
Successes	7.9	3.5	1.60	2.1	982
Persevered	4.4	—	.50	—	—

Computations when all the initial rests are thrown together: .001 × .023 × .010 = .000000230 chance vs. real difference or there are 99,999,997 chances in 100,000,000 of a real difference.

Computations when all the final rests are thrown together: .010 × .014 × .018 = .000000252 chance vs. a real difference or 999,999,748 chances in 1,000,000,000 of a real difference.

failing subjects is practically the same during problem solving but the failing subjects do not decrease their winking rate as much as do the successful subjects during the rest periods.

Those who gave up increase their winking rate just before saying they want to quit working and the rapid winking rate has a tendency to persist for some time afterwards. During the rest periods after the problems the subjects who gave up wink more rapidly than those who succeed. This is true in only about nine-tenths of the cases during initial rest.

Subjects who persevere but fail in solving the problems wink definitely more rapidly during their rest periods before and after each problem than do subjects who succeed.

Winking Rate During Other Activities

In this division of our work no attempt was made to distinguish between successes and failures. The purpose was merely to give some tasks which are usually found to be easy and some which are usually found to be difficult and to determine whether the rate of winking is greater during strenuous or during easy mental work.

Five performances were used: 1. Weight discrimination (comparing a series of weights, 92–112 grams, 4-gram intervals, with a 100-gram standard).

2. Checking the weights by turning them over and looking at labels.

3. Mental multiplication (23 × 244).

4. Repeating alphabet forward twice.

5. Repeating alphabet backward once.

Before and after each of these activities there was a rest period. The rest period after one activity served also as the rest period before the next activity. The tasks were given in chance order.

The inter-wink intervals were longer during the checking of the weights than during the rest periods just prior to and just after this activity. This is because the subjects used their eyes in looking at the weights, and hence, winking rate was decreased.

TABLE IX

EXPERIMENTAL FINDINGS FOR EYE-WINK EXPERIMENT FOR OTHER ACTIVITIES

	Initial	Weight Lifting			Weight Checking			Mental Multipli			Alpha. Forward			Alpha. Backward		
		Rest	Lift	Rest	Rest	Check	Rest	Rest	Prob.	Rest	Rest	Alph.	Rest	Rest	Alph.	Rest
N.....	30.	29	29	29	30	30	30	30	30	30	30	30	30	30	30	29
Mean I.-W. interval	7.2	6.3	8.1	6.8	6.0	12.1	7.4	6.9	3.9	6.5	7.4	6.50	6.4	6.8	5.4	5.80
M_1-M_2.....			1.2			3.5		4.4			1.0			1.2		
$\overline{S.D._D}$.....			1.0			2.5				4.5		.14				.33

The top row of reliability quotients is a comparison of the respective first rest period I.-W. intervals with those of the respective activity periods. The quotient is put under the larger average I.-W. intervals in the different cases.

The bottom row of reliability quotients is a series of comparisons of the last rest periods with those of the respective activity periods.

During mental multiplication the inter-wink intervals were decidedly shorter than they were during the rest periods immediately prior to and immediately succeeding the working of the problems. The correlations of +.40 and +.46 indicate that this group finding is fairly likely for the individual members of the group.

In the cases of weight lifting, repeating alphabet forward and repeating alphabet backward, the reliability quotients range from a reversal of .33 to 1.2 which indicates that the rate of winking during these easier mental activities is practically the same as during rest.

Conclusions

1. Successful subjects wink faster when being presented a problem, when solving it and when reporting the solution than they do during rest. They tend to decrease their rate of winking as they approach the end of the solving period. They also tend to wink less often during rest periods before and after the problem situations than do failing subjects, but the winking rate during mental work in the case of successful and failing subjects is not materially different.

2. Whether successful or not, the subjects tend to wink more rapidly during the presentation of a problem than they do during rest. However, this is statistically certain for successful subjects, but is not statistically certain in case of those who fail.

3. Subjects who give up decrease their winking rate in the third quarter of the solving period and then tend to increase very slightly this rate just prior to informing the experimenter that they cannot solve the problem:

4. Subjects who persevere but fail to solve the problem successfully tend to wink at a constant rate during the whole solving period. They also tend to decrease their winking rate during attempted mental work less than do subjects who give up or subjects who succeed.

5. Winking is more rapid during difficult mental multiplication than it is during such exercises as judging weights or repeating the alphabet forward or repeating the alphabet backward.

Summary

Increase in winking rate accompanies successful problem solving. Such an increase is more commonly present during difficult mental tasks than during easier tasks. Successful problem solving is followed by definite decrease in winking rate while the decrease in winking rate is not so definite in the case of those who fail. Upon analysis, it is found that there is not a reliable difference in the rate of winking during problem solving for successful and failing subjects. The real difference is found when the winking rate during rest is compared with the rate during mental effort.

The hypothesis made by Ponder and Kennedy that the eye-wink is a relief mechanism is not substantiated by our findings. Those investigators based their statement upon observations of people in the library and it is very probable that reading is an entirely different type of mental performance and requires a different type of activity than does problem solving.

Our investigation does not support the drainage theory which states that when one type of work is done that the energy is drained away so that another kind of work cannot be carried on. Rather, we find that there is an increase in winking rate during successful problem solving as compared with periods of rest and it can be inferred that (1) problem solving of the type investigated in this research is accompanied by an increase in somatic activity, or (2) problem solving is in whole or in part a type of somatic activity.

Bibliography

1. Bills, Arthur G. The influence of muscular tension on efficiency of mental work. *Amer. Jour. Psych.*, 1927, **38**, 227–51.
2. Book, W. F. *The psychology of skill.* Gregg Publishing Co., New York, 1925.
3. Clark, Ruth. An experimental study of silent thinking. *Archives of Psychology*, No. 48, 1922. Columbia University Contributions to Philosophy and Psychology, New York.
4. Duffy, Elizabeth. The measurement of muscular tension as a technique for the study of emotional tendencies. *Amer. Jour. Psych.*, 1932, **44**, 146–62.
5. Ford, Adelbert. Attention automatization. An investigation of the transitional nature of mind. *Amer. Jour. Psych.*, 1929, **41**, 1–32.
6. Freeman, G. L. Muscular activity and the mental processes. *Psych. Review*, 1931, **38**, 428–49.

7. —— AND KATZOFF, E. F. Muscular tensions and instability. *Amer. Jour. Psych.*, Jan., 1932, **44**, 789–92.
8. JACOBSON, EDMUND. Differential relation during reading, writing, and other activities as tested by knee-jerk. *Amer. Jour. Physiol.*, 1923, 675–93.
9. ——. Electrical measurement of neuromuscular states during imagination of movement involving skeletal muscle. *Amer. Jour. Physiol.*, Dec.–Jan., 1929–30, **91**, 567–607.
10. ——. Electrical measurements of neuromuscular states during imagination and recollection of various muscular acts. *Amer. Jour. Physiol.*, July–Sept., 1930, **94**, 22–34.
11. ——. Electrical measurements of neuromuscular states during visual imagination and recollection. *Amer. Jour. Physiol.*, Oct.–Dec., 1930, **95**.
12. ——. Electrical measurements of neuromuscular states during mental activity—Contraction of specific muscles during imagination. *Amer. Jour. Physiol.*, Oct.–Dec., 1930, **95**, 703–12.
13. ——. Electrical measurements of neuromuscular states during mental activity—Variation of specific muscles contracting during imagination. *Amer. Jour. Physiol.*, 1931, **96**, 115–21.
14. ——. Electrical measurements of neuromuscular states during mental activity—A note on mental activities concerning an amputated limb. *Amer. Jour. Physiol.*, 1931, **96**, 122–25.
15. ——. Electrical measurements of neuromuscular states during imagination, recollection, and abstract thinking involving the speech musculature. *Amer. Jour. Physiol.*, 1931, **97**, 200–09.
16. JONES, E. E. The influence of bodily posture on mental activities. *Arch. of Psych.*, No. 6. The Science Press, New York, 1907.
17. LOMBARD, WARREN P. Variations of the normal knee-jerk and their relation to the activity of the central nervous system. *Amer. Jour. Psych.*, 1887, **1**, 2–71.
18. LURIA, ALEXANDER. *The nature of human conflict*, tr. by Horsley Gant. Liveright, Inc. New York, 1932.
19. MORGAN, J. J. B. The overcoming of distraction and other resistances. *Arch. of Psych.*, No. 35. The Science Press, New York, 1916.
20. PONDER, ERIC, AND KENNEDY, W. P. On the act of blinking. *Quar. Jour. of Physiol.*, 1927–28, **18**, 89–110.
21. THORSON, AGNES. Tongue movements and internal speech. J. EXPER. PSYCH., 1925, **8**, 1–27.
22. WOODWORTH, R. S. *Psychology*. Revised ed., 1931. Henry Holt & Co., New York.

THE EFFECT OF A CHANGE OF BACKGROUND ON RECALL AND RELEARNING[1]

BY STANLEY G. DULSKY

I. Problem and Introduction

This experiment was designed to answer two questions: (1) How are the recall and relearning scores affected when retention is measured in an environmental situation different from that which obtained during learning, and (2) What is the relative validity of the two proposed explanations of this effect?

Two general explanations have been proposed to account for the deleterious effect. The first assumes that, in learning, the material becomes associated with some feature or features of the background and that it is the absence of these cues which causes the poorer scores made in testing retention. The second explanation assumes that the absence of the old cues exerts no effect, but that the new context or environment, by disturbing the subject, causes the deleterious effect on recall and relearning. Previous investigators, due to the lack of an appropriate procedure, have offered no way of decisively estimating the relative efficacy of these explanations. This experiment was designed to vary the conditions in such a manner that their efficacy could be tested. By our procedure we can determine whether both explanations are correct, and if so we can judge the relative effectiveness of each.

It is a well-known fact that recall depends on environmental conditions. Carr (1) states: "An experience can be recalled most readily in those environmental situations with which it has the most direct, the strongest, and the most

[1] From the Psychological Laboratory of the University of Chicago.

The writer wishes to express his gratitude to Mr. Maurice H. Krout of Crane Junior College, Chicago, and to Mr. David P. Boder of Lewis Institute, Chicago, for their aid in facilitating the experimentation in their respective schools.

numerous associations." The investigations which have studied the effect of a change of background on retention may be divided into three classes on the basis of the method of measuring retention.

The first group of studies measured retention by recall. Pan (**6**) had subjects learn twenty pairs of logically unrelated words of common usage by the paired associate method. The material was partially learned and recall was tested after twenty-four hours by a written reproduction. In a subsidiary experiment, twenty-four pictures of human faces were learned with corresponding names. Retention was tested after forty-eight hours. Learning and recall took place under six different conditions: (1) no context present, (2) a context of logically unrelated words, (3) words logically related to the response words, (4) words logically related to the stimulus words, (5) words logically related to stimulus and response words, and (6) a context of numbers. The removal of the number-context and removal of word context logically related to the response words were detrimental to recall.

Reed (**9**) investigated the effects of three types of change on recall: changes in posture, in mode of stimulation, and in form of response. Changes of posture (sitting and standing) had no effect on recall. Changes in mode of stimulation (auditory and visual) were detrimental to recall. Changes in the form of response (oral and written) gave a smaller percentage of high and perfect scores in recall than the control groups.

A second group of studies measured retention by the method of relearning. Smith and Guthrie (**10**) report two experiments performed in their laboratory by Mr. W. R. Wilson. In the first experiment, ten subjects learned nonsense syllables and relearned them seventy-two hours later. There were four conditions in the experiment: (1) Material learned and relearned in the laboratory; (2) learn in laboratory and relearn out of doors; (3) learn out of doors, relearn out of doors; (4) learn out of doors, relearn in the laboratory. In eight of the ten subjects there was a greater saving in each case where relearning occurred in the same surroundings in

which the first learning had taken place. An average of 11.4 percent fewer repetitions were required to relearn in the same surroundings.

In a second experiment the subject was seated in a room and copied a list of letters on a typewriter whose keys had been arranged in a random order. He copied the same list twenty-four hours later and the times were compared. A variable condition consisting of the presence or absence of the odor of oil of peppermint was used. The procedure was similar to the method used in the first experiment. The average saving in the time required to repeat the list was 9.4 percent greater when the repetition took place under the condition of the first practice.

There has been one experiment, recently published by Pessin (7), that failed to confirm the findings of earlier investigators. He had thirty-two subjects learn lists of seven nonsense syllables by the anticipation method until they could give one perfect reproduction. Learning took place under two conditions: (1) combined auditory and visual stimulation, and (2) quiet conditions. The group that learned in condition (1) relearned the same way and also relearned in condition (2). The group that learned in condition (2) relearned the same way and also relearned in condition (1). Relearning took place after four minutes.

Pessin secured no reliable differences in the data, though there was a slight tendency for the relearning scores to be better in the dissimilar conditions. He states on page 433, "The results obtained in this experiment show that individuals are able to relearn material under dissimilar conditions at least as efficiently as under similar conditions." His explanation is that the individuals increased their efforts to master the material.

It is probably true that the motivation factor accounted for the good showing made by the subjects under the dissimilar conditions. Pessin states that "individuals are able to relearn material under dissimilar conditions at least as efficiently as under similar conditions." This relearning may be done "as efficiently," judged from the standpoint of

errors and repetitions, but, perhaps, not as efficiently judged from the standpoint of the subject's output of energy. Morgan (**5**) showed that individuals could overcome poor working conditions and distractions by increasing the energy output. It is the opinion of the present writer that had motivation been constant in all of Pessin's conditions the results would have been in accord with those secured by earlier investigators of this problem.

A third group of studies has been carried on with animals as subjects. These studies antedated those performed with human subjects. The animals had to reinstate skilled acts in a different environment from that in which they had been learned. Experiments by Watson (**11**), Porter (**8**), Hunter (**4**), and Carr (**2**) have shown that a change of environmental conditions is usually deleterious to the reinstatement of a skilled act. For a summary of these studies see Pan (**6**).

While the results of the investigations agree, mainly, that a change of background may be detrimental to retention, there is no way of decisively judging the efficacy of the two explanations. Pan's experiment supports the association conception. He had a distinctive environment or context for each pair of words. The removal of the word context logically related to the response words and the removal of the number-context were detrimental to recall. A context of logically unrelated words had no appreciable effect on recall when it was removed during the measurement of retention. Thus it would appear that the absence of cues caused the poorer scores made in recall and not the introduction of the new context.

In the experiments by Reed, Wilson, and Pessin, the environment is the same for the entire list of material. There is thus but little possibility for each of the items to become associated with a distinctive aspect of the environment. Here, presumably, we are dealing with the influence of a new context which disturbs the subject and causes the deleterious effects on recall.

In the animal experiments there is but one response or series of responses to the situation and these responses could

get associated with various features of the environment so that these environmental features would be used as cues to the proper responses. The changed environment may eliminate some of these cues. It is also possible that the new environment or background might have disturbed the animals and thus caused the poorer scores in recall. These experiments are indecisive in estimating the relative efficacy of the two factors.

The present investigation corroborates the general finding that a change of background may be detrimental to retention when measured by recall and relearning and, further, throws light on the possible causes.

II. Material and Procedure

A. Learning.—The material to be learned consisted of ten pairs of nonsense syllables which were to be memorized by the paired associate method. All of the syllables were selected from the zero association value list compiled by Glaze (3). Each pair of syllables was printed on a stiff card three by five inches. There were eight sets of cards, ten in a set, used in learning. Four of the sets had a pair of syllables mounted on each card; four of the sets had the stimulus syllable only. This latter group was used to test the learner's progress after each learning trial. The backs of the cards were numbered in such a manner that the experimenter, after memorizing the response and number, could ascertain when the subject gave the correct response.

Distinctive backgrounds were secured for each pair of syllables by means of variations in chroma and differences in brightness. The chroma and grey were secured by means of colored papers which were pasted on the cards.[2] As the card was white, nothing had to be done to secure this degree of brightness. The distinctive backgrounds, then, were formed by using nine chromas, a grey, and a white.

Two types of backgrounds were employed in this study: homogeneous or non-distinctive, and heterogeneous or distinctive. The homogeneous background was identical for a series of cards, that is, it was either all grey or all one color. The heterogeneous background was distinctive for each card or part of card within a series. This type of background was formed by the nine chromas and white. This distinction between the two types of backgrounds will be clarified later.

This study consisted of four experiments. In Experiment I, the left half of each card, the stimulus background, was distinctive. The right half, or response background, was a homogeneous grey. Each card in the set had a different chromatic

---

[2] These were Munsell poster papers secured from Favor, Ruhl and Co., Chicago. Following is a list of the colored papers with their identification numbers.

Light Red—1	Light Blue—25
Maximum Red—4	Maximum Blue—28
Maximum Yellow-Red—8	Middle Purple—34
Maximum Yellow—12	Maximum Purple—36
Maximum Green—20	Middle Grey—42.

stimulus background. For example, one card was half red and half grey; another card was half green and half grey; and so on for each of the ten cards. Experiment II was similar to Experiment I except that the stimulus side was homogeneous or grey and the response side heterogeneous or distinctive. Thus, one card in this experiment was grey on the left side and green on the right; and so on for each of the ten cards. Experiment III had a common chromatic background for both stimulus and response but a different background for each of the ten cards. For example, one pair of syllables was mounted on a red card, another pair on a green card, and so on for each of the ten cards in the experiment. Experiment IV was similar to Experiment III in that there was a common background for both stimulus and response but unlike in that the chromatic backgrounds for the various pairs of syllables were identical.

The pack of ten cards was placed in the proper receptacle of a black metal exposure apparatus. The first card in each set was a blank one. The experimenter then pulled out the last card, *i.e.*, the one nearest him, and slid it in front of the blank one. This procedure was continued with the other cards until the blank one was again in front facing the subject. Each card was exposed for five seconds. At the end of each trial the entire pack of cards was removed and shuffled. This occupied ten seconds. Each trial was a minute apart. Three learning trials were given before the subject was tested for recall. Learning and recall trials were then alternated until the subject could give ten correct responses in one trial. During the learning the experimenter recorded the number of syllables given correctly each trial and the number of trials necessary to learn.

The subject and experimenter were seated opposite each other at a table. The exposure apparatus faced the subject. It was built in such a manner that the subject could not see the experimenter manipulate the cards or record the responses. The following typewritten instructions were handed to the subject:

> You will be shown a series of 10 cards. On each card is printed a pair of syllables, one on the left side and one on the right. You are to learn these two syllables in such a manner that when only the syllable on the left side is exposed you will be able to spell aloud the syllable on the right side of that card.
>
> I will give you three learning trials and then a test for recall. After that there will be an alternation of learning and recall trials until you have learned the entire series.

The experimenter made certain that the directions were fully understood before beginning. Nothing was said about chromatic backgrounds or about a test of retention after the learning had been completed.

B. Recall and Relearning.—Immediately after learning, the subject was given a magazine and told to look through the advertisements until the experimenter stopped him and asked which advertisement he considered the best. After four minutes and forty-five seconds, the subject was told to stop and select the best advertisement. Five minutes after the completion of learning the subject was ready for the recall test. If the recall score were not perfect, relearning was given until the subject could again give the ten responses correctly. During recall and relearning the experimenter recorded the exact response of the subject for each stimulus.

There were three conditions of recall in each of the first three experiments. Condition A of each experiment was the control condition. The subject recalled and relearned the material with the same distinctive background that was present in learning. Under condition B the subject learned as in condition A, but recalled and relearned with the distinctive backgrounds interchanged. For example, if a subject

were in Experiment III, he learned the pair, TOV-JEX, on a red background, and the pair, QIF-GAH, on a green background. When tested for recall, TOV would be shown on a green background and QIF on a red one. Hereafter, condition B will be referred to as the *interchanged* condition. Under condition C the subject learned as in condition A but recalled and relearned with the entire card presenting a grey background. This condition will be referred to as the *neutral* condition. Experiment IV included only conditions A and C.

Eight sets of cards were used. Six sets for the three B conditions, and two sets for the C conditions. One of the two sets in each of the three B conditions had the pair of syllables on each card; the other set had the stimulus syllable only. One of the two sets in the C conditions had the pair of syllables on each card; the other set had the stimulus syllable only.

Table I shows the procedure followed for all the conditions in the four experiments.

TABLE I

Experiment	Learning		Recall and Relearning	
	Stimulus Side	Response Side	Stimulus Side	Response Side
I-A	Chromatic heterogeneous	Grey	Chromatic heterogeneous	Grey
I-B	Chromatic heterogeneous	Grey	Chromatic heterogeneous (Interchanged)	Grey
I-C	Chromatic heterogeneous	Grey	Grey	Grey
II-A	Grey	Chromatic heterogeneous	Grey	Chromatic heterogeneous
II-B	Grey	Chromatic heterogeneous	Grey	Chromatic heterogeneous (Interchanged)
II-C	Grey	Chromatic heterogeneous	Grey	Grey
III-A	Chromatic heterogeneous		Chromatic heterogeneous	
III-B	Chromatic heterogeneous		Chromatic heterogeneous (Interchanged)	
III-C	Chromatic heterogeneous		Grey	
IV-A	Chromatic homogeneous		Chromatic homogeneous	
IV-C	Chromatic homogeneous		Grey	

For example, in the column, 'Learning,' it can be seen that for Experiment I the stimulus side (background) was chromatic heterogeneous. This means that the stimulus sides were colored and that each card had a different color. The response side was homogeneous grey. For recall and relearning in condition A the set-up was the same as for learning. For condition B the learning was the same as for condition A, but the colored backgrounds were interchanged in recall and relearning. For condition C, learning was the same as for condition A, but recall and relearning took place with the syllable backgrounds a homogeneous grey. The procedures for the remaining conditions are read similarly.

C. Subjects.—One hundred ten complete records were secured from approximately 175 subjects. There were thirty subjects in each of the first three experiments and twenty in the fourth. There were ten subjects in each of the eleven conditions—five men and five women in each group. The subjects were all psychology students from the University of Chicago, Lewis Institute, and Crane Junior College.

III. Results

A

What is the effect of a change of background upon retention as measured by recall and relearning scores? This question is answered by a comparison of the data of conditions A and C for all four experiments. Table II gives the total

TABLE II

Condition	Total Number Recalled	Number Recalled Correctly	Number Trials to Relearn
A	391	365	4.8
C	318	265	13.7

number of syllables recalled, the total number recalled correctly, and the number of trials to relearn for the A and C conditions.

The number of syllables recalled was reduced by 73, or 19 percent. The number of syllables recalled correctly was reduced 100, or 28 percent. The number of trials necessary to relearn increased from 4.8 in the A conditions to 13.7 in the C conditions, an increase of 8.9, or 185 percent, over A. In the A condition the percentage recalled correctly of the total number recalled was 93 percent. In the C condition the percentage recalled correctly of the total number recalled was 83 percent. This reduction in correct recall is not an incidental result of the decrease of total recall, for the change affects correctness or accuracy more than it does total recall.

There are two types of incorrectness of recall: (1) The syllable is a part of the response list but is recalled with the wrong stimulus term; (2) The syllable recalled does not belong to the memorized list of responses. Table III shows the

TABLE III

Condition	Incorrect on Basis of Stimulus	Not in Response List
A	11	15
C	43	10

number of syllables in these two categories for the A and C conditions. The change of background materially increased the first class of error. The number was increased from 11 to 43 in spite of the fact that the total recall was decreased. In terms of percentage of total recall the change of background increased this type of error from 3 percent to 13.5 percent.

The effect upon the second type of error is negligible. The change reduced the total number from 15 to 10, but this change is proportional to the reduction of total recall. In terms of percentage of total recall the reduction was from 3.8 to 3.1.

To summarize from the data here presented on the A and C conditions, we may conclude that:

1. A change of background decreased the total number of syllables recalled.
2. A change of background decreased the number of syllables recalled correctly.
3. The reduction in accuracy was greater than the reduction in number recalled.
4. A change of background increased the number of trials necessary to relearn.
5. A change of background increased the number of errors of type 1, *i.e.*, when the syllable evoked is a part of the response list but is recalled with the wrong stimulus term.

B

Does the detrimental effect of a change of background vary with the position of the background? The first three experiments differed only in respect to the position of the background. The distinctive coloring formed the background for the stimulus term in Experiment I, for the response term in Experiment II, and for both the stimulus and response items in Experiment III. The question is answered

by comparing the detrimental effects in the first three experiments. Table IV shows these data.

TABLE IV

Experiment	Total Number of Syllables Recalled	Number Recalled Correctly	Percent Recalled Correctly	Number Trials to Relearn
I-A	97	92	95	0.9
I-C	80	66	83	2.4
II-A	100	97	97	0.2
II-C	68	46	68	6.2
III-A	97	84	87	2.4
III-C	79	68	86	3.4
IV-A	97	92	95	1.3
IV-C	91	85	93	1.7

a. Relative Effect of Stimulus and Response Backgrounds.—The data for this comparison will be found in Table IV under Experiments I and II. A change of the response background (Experiment II) gave the greater reduction in total recall both in absolute and relative terms, a greater reduction in number correct in absolute and relative terms, and a greater increase in number of trials necessary to relearn. We may conclude that a change of the response background is more effective upon retention as measured by recall and relearning than a change in stimulus background.

b. Will a Change of Both Stimulus and Response Backgrounds Give a Greater Detrimental Effect than Either One Alone?—The detrimental effects of Experiment III differ in no consistent way from those of Experiment I. Combining the two changes thus gives about the same effect as that of the less effective component.

C

Will the results of a change from a heterogeneous chromatic background to a homogeneous achromatic background (Experiment III) differ from a change of a homogeneous chromatic background to a homogeneous achromatic background (Experiment IV)? These data can be found in Table IV.

The results of Experiment IV show a consistent, but statistically unreliable, difference by all criteria in favor of the control condition. As compared with the results of Experiment III, this type of change is less effective. We may conclude, then, that a change from a heterogeneous chromatic background to a homogeneous achromatic background is more detrimental to retention than a change from a homogeneous chromatic background to a homogeneous achromatic one.

D

Will the response term become associated with the distinctive background during learning and hence the absence of this background at recall presumably diminish the amount and correctness of recall?

Condition B was introduced to test the validity of the first hypothesis, viz., that each response item becomes associated not only with the stimulus term but also with the distinctive backgrounds in such a way that these backgrounds will tend to reinstate the response, and hence the absence of these backgrounds at the time of recall will decrease the chances of a successful recall. If the response term is not associated with the background, but only with the stimulus term, then a removal of that background and the introduction of a new one (condition B) should give practically the same results in recall as condition C. If this is the case, then the explanation is that the poorer scores in recall must be due to the effect of the novel background.

However, if each response item is associated with its background as well as with the stimulus term, then, in the interchanged condition, the stimulus will tend to arouse one response and the background another. In this situation, some of the stimulus terms may arouse the correct responses and some of the backgrounds may evoke their appropriate responses depending on the relative strength of the associations. If the two are of equal strength, the resulting conflict may bring no response, or some unusual type of response. Therefore, we recorded the response given to each stimulus in the recall test and classified them as to whether they were:

(1) correct in respect to the stimulus term; (2) correct in respect to the background; or (3) incorrect in reference to both. The third classification was subdivided according to whether the responses belonged to the list learned or whether they were novel. These results will enable us to evaluate the efficacy of the first explanation; and any difference in the results between B and C will enable us to determine the validity of the second hypothesis.

Table V shows the data on the B conditions for the first

TABLE V

Experiment	Total Number Syllables Recalled	Number Recalled Correctly on Basis of Stimulus Term	Percent	Number Recalled Correctly on Basis of Color	Percent
I-B.....	72	41	57	24	33
II-B.....	80	37	46	34	43
III-B.....	86	51	60	22	26

three experiments. The data prove that the response item does become associated with both the stimulus syllable and the background. The stimulus syllables evoked 129 correct responses (summation of I, II, III-B), and the background 80. These figures thus represent the relative strength of the two associations when brought into conflict.

The data show that the background in Experiment II aroused a larger number of correct responses (on the basis of color) than the backgrounds in the other two experiments, both in absolute terms and in proportion to the total recall. This proves that the strength of the association in the background is much greater in Experiment II, *i.e.*, when the response background is chromatic or distinctive. This fact is in harmony with the result noted previously, that the change of the response background is more detrimental to recall and relearning than is any other position.

The conflict between stimulus and background at recall materially reduces the total number of items recalled as compared with normal condition A. Table VI shows these data. In the A conditions of the three experiments, 294 syllables were recalled as opposed to 238 in the B conditions.

TABLE VI

Condition	Total Number Syllables Re-called	Number Recalled Correctly on Basis of Stimulus Term	Number Recalled Correctly on Basis of Color	Number Responses in List	Number Responses Not in List	Number Relearning Trials
A	294	273	—	8	13	3.5
B	238	129	80	19	10	15.6

Let us combine the number correctly recalled by the stimulus terms and those correctly recalled by the backgrounds in the B conditions and term this the total correct recall. The conflict reduces this total correct recall, as compared with condition A, quite materially 273 to 209. It also reduces the total correct relative to the total number 93 percent to 87 percent.

The conflict increased the first type of error as compared with condition A in spite of the reduction in the total number recalled. In relative terms the increase was much greater. The conflict reduced the second type of error when stated in absolute terms, but did not when stated in relative terms.

The conflict increased the number of trials necessary to relearn so that the recall would be correct in terms of the stimulus syllable; *i.e.*, on the average it requires a certain number of trials to ignore these associations of the response term with the background.

The experiment proves that the response term does get associated with a distinctive background and that this background evokes the appropriate syllable even when it is brought into conflict with the associative tendency of the stimulus syllable. The presence of this background at the time of recall is thus conducive to accuracy—a greater degree of correct recall.

E. Comparison of the Data of Conditions B and C

We stated earlier in the paper that a comparison of the data of conditions B and C would give us information about the relative validity of the two hypotheses in explaining the deleterious effect of a change of background on retention as

measured by recall and relearning. Table VII gives certain pertinent data on both conditions. The normal condition, A, is included for comparative purposes. These data show that the C conditions were less deleterious to recall and

TABLE VII

Condition	Total Number Syllables Recalled	Number Recalled Correctly	Percent Recalled Correctly	Number Trials to Relearn
A	294	273	93	3.5
B	238	129	54	15.6
C	227	180	79	12.0

relearning than the B conditions. Except for the total number of syllables recalled, each comparison shows a convincing difference. This is convincing proof that lack of cues (condition B) is a more potent factor than novel background (condition C) in producing deleterious effects on retention. Our C condition differs from the usual 'change of background' procedure in experimentation because in the learning each pair of syllables had a distinctive background and these ten distinctive backgrounds were changed to one homogeneous background. In the usual experiment the learning all takes place with one background and then in recall the entire background is changed to another. This type of procedure is exemplified by Experiment IV–C. In this experiment the syllables were learned on an orange background and recalled and relearned on a grey background. Table VIII shows the data of Experiment IV.

TABLE VIII

Condition	Total Number Syllables Recalled	Number Recalled Correctly	Percent Recalled Correctly	Number Trials to Relearn
IV-A.....	97	92	95	1.3
IV-C.....	91	85	93	1.7

It can be seen that the differences are very small, statistically unreliable, and that consequently the explanation of novelty does not hold because here we have a novel back-

ground and no effect on retention as measured by recall and relearning. This supports the statement made earlier, when the data of conditions B and C were contrasted, that lack of cues due to change of background is a more potent factor than novelty—disturbing the subject—in producing deleterious effects on retention.

IV. Conclusions

Three conclusions may be inferred from our data:

1. When the materials become associated with the environment during learning, any change of that environment is likely to prove detrimental to recall. The degree of effect will presumably vary with the number and strength of the connections, and the extent to which the environment is altered. All degrees of effect may be secured.

2. When no such associations are present, the introduction of new environmental conditions may distract the subject and interfere with recall. The degree of interference will presumably vary with the degree and character of the change, and the subject's susceptibility to it.

3. There is thus no need to assume that every environmental change will be detrimental. Some changes may be too slight to produce an effect. The experiments prove that the environmental conditions are one of the factors involved in recall, and any experiments yielding negative results do not disprove the principle.

(Manuscript received August 22, 1934)

References

1. Carr, H., *Psychology*. Longmans, Green, 1925.
2. Carr, H., Maze studies with the white rat. I. Normal animals. *J. Animal Beh.*, 1917, 7, 259–75.
3. Glaze, J. A., The association value of nonsense syllables. *Ped. Sem.*, 1928, 35, 255–69.
4. Hunter, W. S., Some labyrinth habits of the domestic pigeon. *J. Animal Beh.*, 1911, 1, 278–304.
5. Morgan, J. J. B., The overcoming of distraction and other resistances. *Arch. Psychol.*, 1916, No. 35, 84.
6. Pan, S., Influence of context upon learning and recall. J. Exp. Psychol., 1926, 9, 468–91.

7. PESSIN, J., Effect of similar and dissimilar conditions upon learning and relearning. J. EXP. PSYCHOL., 1932, **15**, 427–35.
8. PORTER, J. P., Further study of the English Sparrow and other birds. *Amer. J. Psychol.*, 1906, **17**, 248–271.
9. REED, H. J., The influence of a change of conditions upon the amount recalled. J. EXP. PSYCHOL., 1931, **14**, 632–649.
10. SMITH AND GUTHRIE, *General Psychology*. Appleton, N. Y., 1921. Pp. 112–113.
11. WATSON, J. B., Behavior of noddy and sooty terns. *Carnegie Institute Publ.*, 1909, No. 103.

TIME-ORDER ERRORS IN THE DISCRIMINATION OF SHORT TONAL DURATIONS[1]

BY LELAND H. STOTT

University of Illinois

The purpose of the present investigation was to make an accurate and thorough determination of the constant, or time-order errors (C.E.'s), which occur in comparing short tonal durations (intervals of time filled with continuous tone). These errors, as well as the location of the indifference duration, *i.e.*, the duration at which time-order errors disappear, were studied under a considerable variety of carefully controlled conditions, with a total of five hundred and twenty-four subjects from whom no less than 99,480 judgments were obtained. The durations used in the investigation ranged from 0.2 to 36.0 sec. Most of the work, however, was concentrated upon those between 0.5 and 2.0 sec.

The preceding experience of the subject in comparing durations was revealed by the present study as an enormously important factor in determining the time-order errors and the indifference duration. This experience may consist of previous sittings, or of earlier trials of a given sitting. Previous sittings are particularly influential in the case of a long series of sittings with a wide range of standards. The earlier trials at a given sitting, under certain conditions, also result in a marked change in the constant errors (and consequent shift in the indifference point) shown by later trials; a change which readily comes to light in long sittings when the data are fractionated for successive portions of the sitting.

This factor of experience is one which has been very largely neglected in the past. In the present study it is given credit

[1] This paper is largely an abridgment of a thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Illinois, 1933. The problem was suggested by Prof. Herbert Woodrow, under whose supervision the investigation was carried out, and to whom the writer is deeply indebted.

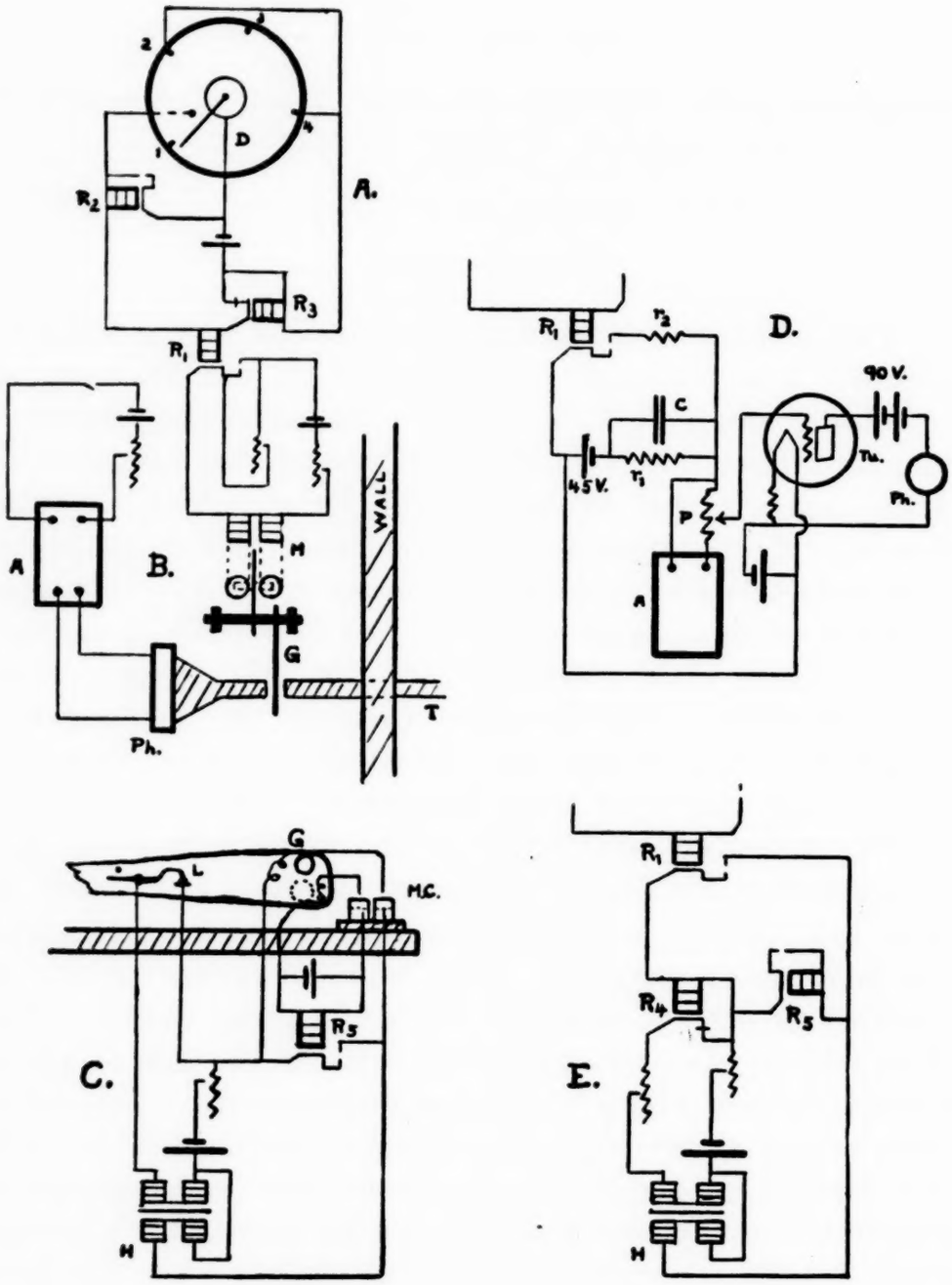

FIG. 1. Diagram of apparatus

A. The time control device.

D, copper disk 12 inches in diameter.

1 and 3, movable copper contact points in direct contact with the disk.

2 and 4, movable copper contact points insulated from the disk.

for having produced some very perplexing situations, and only when it was recognized and adequately controlled was it possible to obtain definite and understandable results.

The investigation consisted of a series of four experiments. The first one, which consumed a total of 350 hours of the time of fifteen subjects, resulted in the establishment of an indifference duration in the neighborhood of 1.67 sec. The constant errors were positive (underestimation of the second duration in comparison with the first) for all durations shorter, and negative for all durations longer than the indifference point. The lengths used covered a wide range, extending from 0.2 to 36.0 sec. The second experiment which covered durations of 0.8 to 2.64 sec. was an attempt to verify the results of the first. The result was the appearance of *four* indifference points, one for each standard, and none of which

---

R_1, relay which always controlled the sound making device.
R_2, shunt relay.
R_3, relay which served to break circuit which activated R_1.

B. The electro-mechanical shutter apparatus for presenting sounds.

M, electric magnets which operated the shutter.
G, aluminum shutter attached to shaft with lever.
T, glass tubes which conducted the sound to subjects' room.
Ph., phone.
A, audio oscillator.

C. Set-up for timing the shutter.

G, shutter.
H, Hipp chronoscope.
M.C., mercury cup contacts.
R_6, relay which served to hold intact the lower chronoscope magnet circuit until the shutter descended and made contact at *M.C.*

D. The electrical set-up for presenting sounds to groups.

Tu., a 112A radio tube.
Ph., phone.
A, audio oscillator.
C, condenser of 0.1 m.f.d. capacity.
r_1 and r_2, resistances of 250,000 ohms each.
P, potentiometer by means of which the loudness of the sound was regulated.

E. Set-up for timing the interval between the make and the break of contact in the time control relay, R_1.

H, Hipp chronoscope.
R_4, relay which effected a break in the upper magnet circuit of the Hipp.
R_5, shunt relay which served to hold intact the lower magnet circuit until R_1 was activated. This relay had to be closed by hand before each time measurement was made.

was similar in length to the first one. In this case twenty-six subjects served eight hours each for a total of 208 hours of experimental time.

It was at this point that the extreme importance of the subject's previous experience in determining the character of his discriminations was discovered. In the work that followed *different* groups of around a hundred naive subjects were used in short sittings with each standard. Under these conditions durations ranging from 0.5 to 1.92 sec. were investigated and the time-order indifference point, at the initial stage of practice, was definitely established at about 0.92 sec. for tone filled intervals. The first indifference point of 1.67 sec., then, must be considered to be largely a resultant of the great amount of experience obtained by the relatively few subjects during the long series of prolonged sittings involving hours of listening to very long durations. The variety of indifference points obtained in the second experiment is likewise charged to experience, but particularly to the effect exerted on the results of any one whole sitting of the earlier trials in that same sitting.

These four experiments together with their results will be discussed in detail in the order in which they were performed. First, however, a detailed description of the apparatus and the modifications which were made in it as the investigation progressed will be given.

Apparatus

For an investigation of this nature rather complicated apparatus was required in order to provide for the necessary accuracy in presenting and timing the sounds. In all instances the set-up consisted of three main divisions. These were (1) a time control device, (2) an apparatus for producing and presenting the sound stimuli, and (3) a set-up for measuring the durations of the sounds presented.

Figure 1, *A*, shows the essential features of the time control device. It consisted of a brass disk (*D*) twelve inches in diameter mounted on a heavy tripod in a horizontal position. The disk was graduated by machine about its circumference in degrees. A Telechron synchronous motor was mounted directly over the center of the disk, but not in contact with it, in such a manner that when the motor revolved an arm which was attached to its shaft swung around above the edge of the disk. Two pairs of movable brass contacts were fitted to the edge of the disk. These could be accurately set at any point on the edge of the disk by means of set screws. One of each pair of contacts (1 or 3) was in direct connection with the disk while the other one (2 or 4) was insulated carefully from it. A platinum wire brush attached to the end

of the motor arm connected in turn with each of the four contacts as it revolved. A six volt battery and two pony relays (R_1 and R_2), 20 ohms resistance each, were connected in circuit with the body of the motor and the disk so that when the brush of the motor arm made contact at 1 the circuit was completed and R_1 and R_2 were activated. R_1 served to close the sound circuit while R_2 formed a short circuit through its own coils. In that way the disk was cut out of the circuit and connection was maintained at R_1 after the momentary connection at contact 1. A second circuit was completed when the motor arm made contact at 2 or 4. This circuit momentarily activated a third relay (R_3) which served to break the first circuit and thus to release R_1 and R_2 breaking connection in the sound circuit.

At the beginning of Experiment II a minor change was introduced in this device which was quite important from the standpoint of accuracy and preciseness. In order to eliminate the necessity of resetting the contact points on the control disk for each stimulus presentation a switch-board was incorporated in the set-up. With this arrangement fifteen contact points, instead of the original three movable ones, were used. Before the work on any particular standard was begun each contact was carefully set in one of the fifteen positions necessary for the presentation of the various combinations of standard and variable. Careful measurements were taken by means of the Hipp chronoscope for each contact whose position was changed slightly back and forth until the average of ten measurements was within 1 σ of the intended duration. After the contacts were once set they remained in exactly the same positions for all the work on that standard. These contacts were insulated from the disk. By means of three rows of five jacks on the switchboard, each connected by wire to one of the fifteen contact points, and three plugs corresponding to the three original movable contacts any desired stimulus combination was quickly and easily obtained. The inaccuracies of contact setting were thus eliminated.

Since our object was to study the discrimination of sound durations it seemed important that these sounds be smoothly continuous and as free from initial and final clicks and other extraneous noises as possible. At the same time it was necessary that their beginnings and endings be sharp and definite. A means must also be provided for the accurate measurement of their objective durations. The problem of meeting these criteria was not simple. A great deal of preliminary work was required before an entirely satisfactory set-up was found. Two devices were actually used. The first (Fig. 1, *B*) consisted essentially of an electro-magnetically operated gate or shutter (*G*) which controlled the passage of a continuous sound through a glass tube (*T*), thence to the subject's ears via rubber tubing and a dictophone head-set. The continuous sound was produced by a radio ear phone (*Ph.*) connected to the out-put of an audio oscillator (*A*) of a frequency of 1000 d.v. The tone emitted at the head-set was clear and even and of the proper intensity, and was relatively free from clicks and noises. This device proved, on the whole, to be fairly satisfactory for use with individual subjects. Its most serious defect was discovered when actual duration measurements of the sounds which it presented were made.

Several methods of timing the movements of the time control shutter were tried out. The one which proved mosts atisfactory is diagrammed in Fig. 1, *C*. A Hipp chronoscope (*H*) was connected through mercury cups (*M*) and a lever contact (*L*) with the shutter in such a manner that direct timing of the movements of the shutter was possible. As the shutter was drawn down to open the sound passage way a break in the upper chronoscope magnet circuit was effected at *L*. This started the hands of the chronoscope. As the shutter returned to its original position the second chrono-

scope circuit, which had been held intact by the shunt relay (R_5) before the shutter descended, was broken and the hands of the chronoscope were stopped.

By carefully adjusting the heights of the contacts accurate measurements of the intervals between the different positions of the shutter were made and a good check-up on its action time was obtained. It was found, for example, that the shutter actually required a period of 53.7 σ on the average to pass from the closed position to the wide-open position, and a similarly long interval to return from the wide-open to the closed position. As has been said, in order for the sound stimuli to be entirely satisfactory they must, in addition to being smooth and free from extraneous noises, have a sharp and definite beginning and ending. Even though these short periods of building up and dying down as measured were not noticeable to the subject they were nevertheless undesirable since they rendered an exact description of the durations very difficult. These measurements indicated, however, that when the half-open position of the shutter was arbitrarily taken as the beginning and end of the sound interval it measured almost exactly as intended.[2] The durations were thus defined as beginning and ending at the half-open position of the shutter.

The electro-magnetic shutter apparatus, in addition to the defect already mentioned, was obviously not adaptable for use with groups. It was therefore discarded in favor of the electrical device [3] shown in Fig. 1, *D*. This device was found to be quite satisfactory in every way. It consisted essentially of a radio vacuum tube and a system of resistances and condensers in connection with the audio oscillator. A radio phone, which produced the sound in the subjects' room, was connected to the negative pole of a 90 volt radio 'B' battery. The positive terminal was connected to the plate of a 112A vacuum tube and the return wire from the phone was connected with the negative side of the tube filament. In that way a 90 volt potential difference was produced between the plate and the filament. The grid of the tube was connected to the negative pole of another 45 volt battery through a resistance (r_1) and a potentiometer (P). The potentiometer controlled the input from the audio oscillator. The positive side of the filament was connected with the positive pole of the battery. This gave the grid a potential of about -40 volts with the alternating input of the oscillator imposed upon it. As long as this condition was maintained no flow of electrons could pass between the filament and the plate since the negative charge on the grid repelled them, and consequently no sound was made at the phone. In order to get action between the filament and the plate the negative grid potential had to be reduced. For best working conditions in this particular tube the grid potential must be about -4.5 volts. This condition was brought about when connection was made through the contact points of the relay (R_1) completing another circuit through the grid battery. The resistances in that circuit were such that enough current was taken to cut the grid potential to the proper point. Since the 1,000 cycle alternating frequency of the audio oscillator was imposed upon the grid the action between the filament and the plate was interrupted at that frequency producing a 1,000 cycle tone at the phone. The condenser across the first resistance required a certain small amount of time to charge after contact was made at R_1 and this prevented the change in potential on the grid from taking place completely instantaneously. This slight lag in grid

[2] For example, when the contacts were set for a duration of 0.6 sec. the average of 20 obtained measurements was 0.597 sec.

[3] Thanks are due Prof. E. A. Culler for his kind coöperation and for the assistance of his electrical technician in setting up this sound producing device.

potential change served to prevent too sudden action in the diaphram of the receiver and thus eliminated the usual clicks at the beginning and end of the sound. Condensers were also connected around the points of the relay (R_1) and the contacts on the control disk to absorb sparks arising from contact. In that way all clicks and extraneous sounds were eliminated. The intensity of the sound was easily controlled by means of the potentiometer on the oscillator input.

Since the relay (R_1) served to make and break the circuit which controlled the sound at the receiver the time measurements of its action were considered sufficiently accurate for describing the durations of sound. Fig. 1, *E*, shows the essentials of the measuring device used. It consisted of the Hipp chronoscope, a storage battery, two variable resistance coils and two extra relays. A Klopsteg fall apparatus was used to test and regulate the Hipp which was set to measure the interval between a break in the upper magnet circuit and a break in the lower magnet circuit. The connection between the three relays and the chronoscope which are necessary in order to bring about these two breaks simultaneously with the make and break of the sound control relay may be seen in the diagram.

The method of constant stimuli was used throughout this study. A complete stimulus presentation consisted of a standard duration and a variable duration separated by an interval of silence of 1.5 sec. Each pair was preceded by a warning signal given 1.5 sec. before the beginning of the first sound. This warning signal was a single flash of a small electric bulb which was mounted on the table before the subject. In one half of the presentations of every experimental sitting the standard was given *second* preceded by a variable. The order in which the various standard-variable combinations were presented was determined by an irregular chance order schedule.

Each subject was provided with a type written copy of the instructions which was always kept before him during the experimental period. He was instructed to pay close attention to the length of the sounds at each presentation and to try especially to make each judgment independently without taking into account past judgments. He was also asked to refrain from counting, tapping, marking time, or making any sort of rythmical movement while the sounds were being given.

Experiment I

Experiment I consisted of two parts. In the first part five subjects served one at a time for 38 experimental hours each. The standards used were twelve in number and ranged

from 0.2 to 30.0 sec. in length. The variables in each instance were 0 percent, 5 percent, 10 percent, 15 percent and 20 percent longer than the standard. Each subject began with the shortest standard, and at each successive sitting a longer one was used until the maximum duration was reached (ascending series). Then the order was reversed, beginning this time with the longest and proceeding regularly by taking up at each sitting the next shorter until the shortest was reached again (descending series). One experimental hour was sufficient in most cases to complete the set of 100 presentations for each standard. For the four longest standards, however, two and three hours were required to complete a set.

In the second part of the experiment two groups of four and six subjects respectively served for sixteen experimental hours each. Only the eight shorter standards, ranging from 0.2 to 4.0 sec. were used with the groups. The group of six was given the descending series first followed by the ascending series, while with the group of four the same procedure was followed as with the individual subjects.

Members of the staff of the department of psychology, University of Illinois, and students in psychology, both graduate and undergraduate, served as subjects.

Results.—The method used to show time-order errors was to calculate separately for each variable length the percentage of judgments signifying 'variable longer' for each of the orders of presentation. This was done for each of the eight standards which were worked by all fifteen subjects. These percentages together with the percentage differences between the orders (V-S minus S-V) are given in Table I. The differences represent the time-order errors. For standards of 0.2 to 1.0 sec. inclusive the differences are all positive. This is also true of the three smallest variables of the 1.5 sec. standard. For the two longer variables of the 1.5 sec. standard, and for all durations greater than 1.725 sec., the differences are negative with one exception (when the 2.0 sec. durations were compared the average difference was zero).

The results obtained from the five individual subjects

TABLE I

EFFECT OF ORDER OF PRESENTATION UPON PERCENTAGES OF JUDGMENTS INDICATING THE VARIABLE AS LONGER THAN THE STANDARD

The percentage differences between the two orders (V-S minus S-V) represent the time-order errors.

The same 15 subjects each made 100 comparisons at each of the two sittings with each standard.

Standard (Sec.)	Order of Presentation	Percentages 'Variable Longer' Judgments for All Variables				
		0 percent	+5 percent	+10 percent	+15 percent	+20 percent
0.2		(0.2)	(0.21)	(0.22)	(0.23)	(0.24)
	V-S	57.0	68.0	71.0	86.0	83.0
	S-V	43.0	47.0	56.0	69.0	73.0
	Diffs.	+14.0	+21.0	+15.0	+17.0	+10.0
0.4		(0.4)	(0.42)	(0.44)	(0.46)	(0.48)
	V-S	56.0	71.0	78.0	86.0	84.0
	S-V	44.0	54.0	67.0	74.0	82.0
	Diffs.	+12.0	+17.0	+11.0	+12.0	+ 2.0
0.6		(0.6)	(0.63)	(0.66)	(0.69)	(0.72)
	V-S	57.0	70.0	81.0	88.0	89.0
	S-V	43.0	51.0	62.0	75.0	81.0
	Diffs.	+14.0	+19.0	+19.0	+13.0	+ 8.0
0.8		(0.8)	(0.84)	(0.88)	(0.92)	(0.96)
	V-S	62.0	70.0	78.0	87.0	91.0
	S-V	38.0	49.0	68.0	73.0	86.0
	Diffs.	+24.0	+21.0	+10.0	+14.0	+ 5.0
1.0		(1.0)	(1.05)	(1.1)	(1.15)	(1.2)
	V-S	80.0	71.0	77.0	83.0	86.0
	S-V	40.0	56.0	66.0	74.0	84.0
	Diffs.	+20.0	+15.0	+11.0	+ 9.0	+ 2.0
1.5		(1.5)	(1.575)	(1.650)	(1.725)	(1.8)
	V-S	54.0	65.0	73.0	76.0	85.0
	S-V	46.0	57.0	70.0	82.0	86.0
	Diffs.	+ 8.0	+ 8.0	+ 3.0	− 6.0	− 1.0
2.0		(2.0)	(2.1)	(2.2)	(2.3)	(2.4)
	V-S	50.0	62.0	71.0	72.0	76.0
	S-V	50.0	66.0	76.0	89.0	92.0
	Diffs.	0.0	− 4.0	− 5.0	−17.0	−16.0
4.0		(4.0)	(4.2)	(4.4)	(4.6)	(4.8)
	V-S	47.0	53.0	55.0	67.0	68.0
	S-V	53.0	67.0	78.0	81.0	87.0
	Diffs.	− 6.0	−14.0	−23.0	−14.0	−19.0

of the first part of the experiment with the four longest standards indicate strikingly large negative time-order errors for those long durations. The following tabulation shows the average numbers of times the second duration was judged

longer and shorter during the ascending and descending series combined for each standard, after the results for all variables and for both orders (S-V and V-S) have been pooled. Except for time-order errors, the number of judgments 'longer' should equal the number 'shorter.'

Standard	6.0 sec.	10.0 sec.	20.0 sec.	30.0 sec.
Ave. 'L' judgments	98.8	103.2	100.8	99.4
Ave. 'S' judgments	48.4	48.4	50.8	54.6

These figures show, on the average, practically two 'longer' judgments to one 'shorter'; a very marked tendency to overestimate the second duration.

The results indicate, then, that under the conditions of this experiment, where a few subjects sat repeatedly, making large numbers of judgments for each standard length, the general outcome was a relative underestimation of the second member of the pair for all durations from 0.2 up to 1.5 sec. and a relative overestimation of the second for durations longer than 2.0 sec. up to 36.0 sec. The indifference point under these conditions must therefore lie somewhere between 1.5 and 2.0 sec. Its most probable length is about 1.67 sec. These facts are presented in graphic form in Fig. 2, in which curves representing the percentages of 'variable longer' judgments for each order of presentation are shown.

Experiment II

The results of Experiment I seem quite definitely to indicate a range of lengths within which an indifference point must lie. An exact determination of that length, however, was obviously impossible without more data. The object of Experiment II was to explore intensively this region of indifference and, if possible, to definitely establish an average indifference duration. Four standards of 1.0, 1.4, 1.8 and 2.2 sec. were therefore chosen. Five lengths of variables were again used. In this case, however, they were distributed equally as to length on either side of the standard and covered twice the range of relative length. They were, in relation to

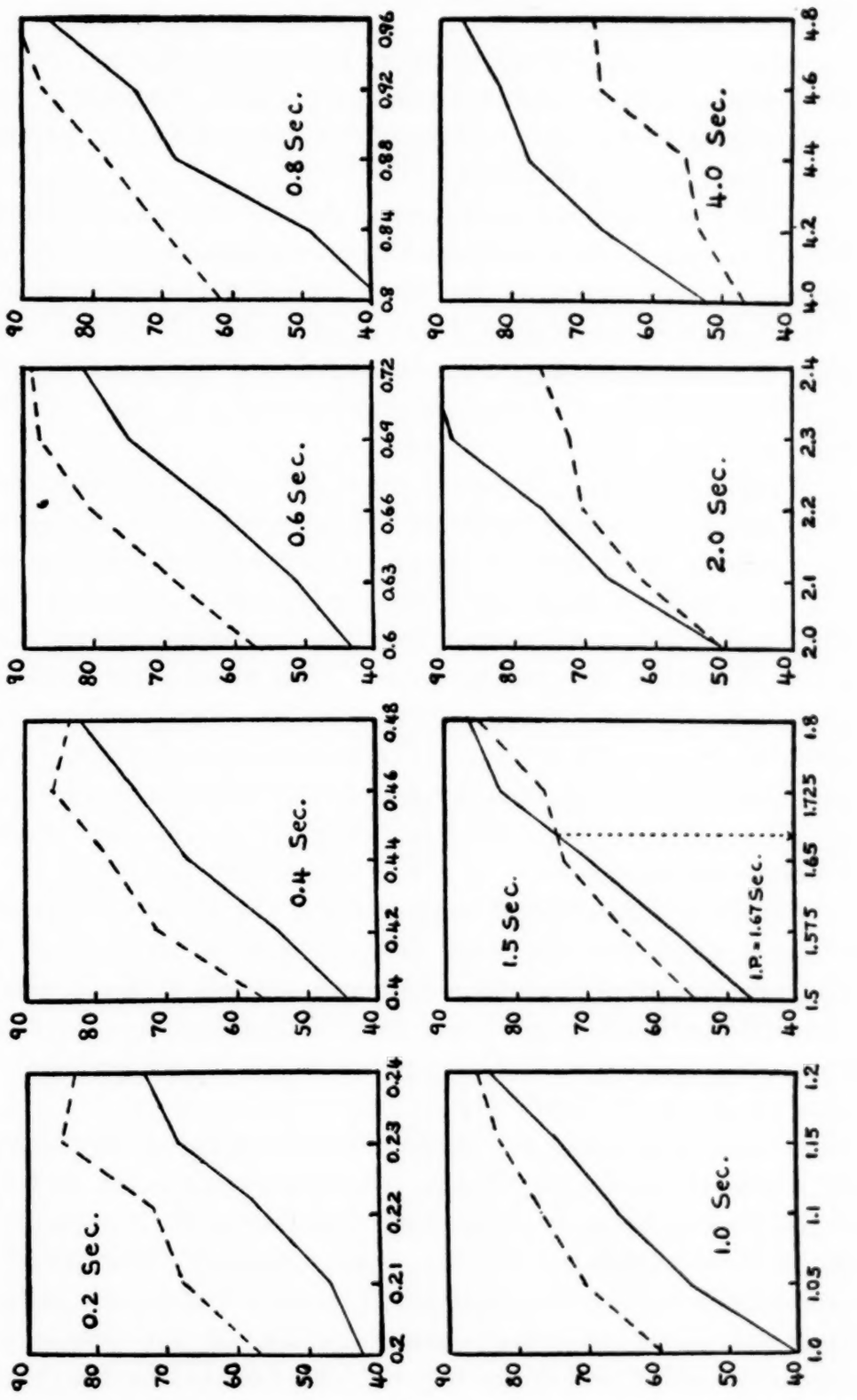

FIG. 2. Graphs illustrating the differences in percentages of judgments which signified 'variable longer' between the two orders of presentation. Data from Table I. Solid lines show the standard first (S-V) order; broken lines show variable first (V-S) order. Each percentage is based upon 300 judgments, 20 from each of 15 subjects. Same subjects served with all eight standards. Variables ranged from 0 percent to 20 percent longer than the standard in each case.

the standard, 20 percent shorter, 10 percent shorter, equal, 10 percent longer and 20 percent longer. Including the variables the total range of durations covered in this experiment was from 0.8 to 2.64 sec.

With the improved tone control device, 150 presentations, fifteen for each standard-variable combination, were given during an experimental period of about fifty-five minutes. As before, only one standard was used on a single day. With exception of the change already mentioned the same experimental conditions were maintained as in the earlier experiment.

Twelve men and fifteen women, students in elementary psychology, served as subjects. They were given the series in groups of from four to six and the experimental schedule was so arranged that all could conveniently complete two sittings a week. All were given the ascending series first followed by the descending series. The twenty-six subjects made a total of 3,900 judgments upon the 150 presentations given them at each sitting. We thus obtained 7,800 judgments for each standard and a grand total of 31,200 judgments for all standards from the ascending series and descending series combined.

Results.—The method used to show the time-order errors in these data was the same as that used in Experiment I. The percentage of the 780 judgments of each variable which signified 'variable longer' was calculated separately for each order of presentation. These percentages for all four standards are shown in Table II.

An examination of this table will reveal a rather perplexing situation in the light of the previous results. It will be noticed that both positive and negative C.E.'s appear in connection with each of the four standards: the shorter variables in each case show positive C.E.'s while the longer variables show negative ones. Instead of the single indifference point which we expected to find in the neighborhood of 1.6 sec. the results show four indifference points ranging between 0.95 and 1.8 sec. in length.

This multiplicity of indifference points, of course, means

TABLE II

EFFECT OF ORDER OF PRESENTATION UPON PERCENTAGES OF JUDGMENTS INDICATING THE VARIABLE AS LONGER THAN THE STANDARD IN THE DATA FROM EXPERIMENT II

The percentage differences between the two orders (V-S minus S-V) represent the constant errors.

The same 26 subjects each made 150 comparisons at each of the two sittings with each standard.

Standard (Sec.)	Order of Presentation	Percentages 'Variable Longer' Judgments for All Variables				
		−20 percent	−10 percent	0 percent	+10 percent	+20 percent
1.0		(0.8)	(0.9)	(1.0)	(1.1)	(1.2)
	V-S	22.37	33.01	47.66	64.49	76.60
	S-V	10.26	28.46	52.34	69.03	78.97
	Diffs.	+12.11	+4.55	− 4.68	− 4.54	− 2.37
1.4		(1.1)	(1.26)	(1.4)	(1.54)	(1.68)
	V-S	19.23	33.59	48.43	60.64	72.12
	S-V	12.03	29.74	51.57	69.81	82.18
	Diffs.	+ 7.20	+3.85	− 3.14	− 9.17	−10.06
1.8		(1.44)	(1.62)	(1.8)	(1.98)	(2.16)
	V-S	18.59	35.26	47.85	62.37	71.28
	S-V	12.12	29.89	52.15	73.53	84.81
	Diffs.	+ 6.47	+5.37	− 4.30	−11.16	−13.53
2.2		(1.76)	(1.98)	(2.2)	(2.42)	(2.64)
	V-S	17.82	30.64	44.04	56.34	70.51
	S-V	15.70	32.69	55.96	78.33	84.49
	Diffs.	+ 2.12	−2.05	−11.92	−21.99	−13.98

that in many instances where the same or similar durations were used in the two experiments C.E.'s of opposite sign appeared in connection with them. This is especially noticable in the case of the 1.0 sec. standard. In Experiment I all C.E.'s were positive with the 1.0 sec. standard. In the present experiment where the same or similar variable lengths (1.0, 1.1 and 1.2 sec.) were involved, the C.E.'s were all negative. These facts may be more clearly seen in the graphs of Fig. 3 which were plotted directly from the data as presented in Table II. The upper halves of these curves, involving the variables greater than the standard in each case, may be compared directly with the curves for similar durations shown in Fig. 2.

The problem of accounting for the appearance of these four new indifference points and the consequent contradiction,

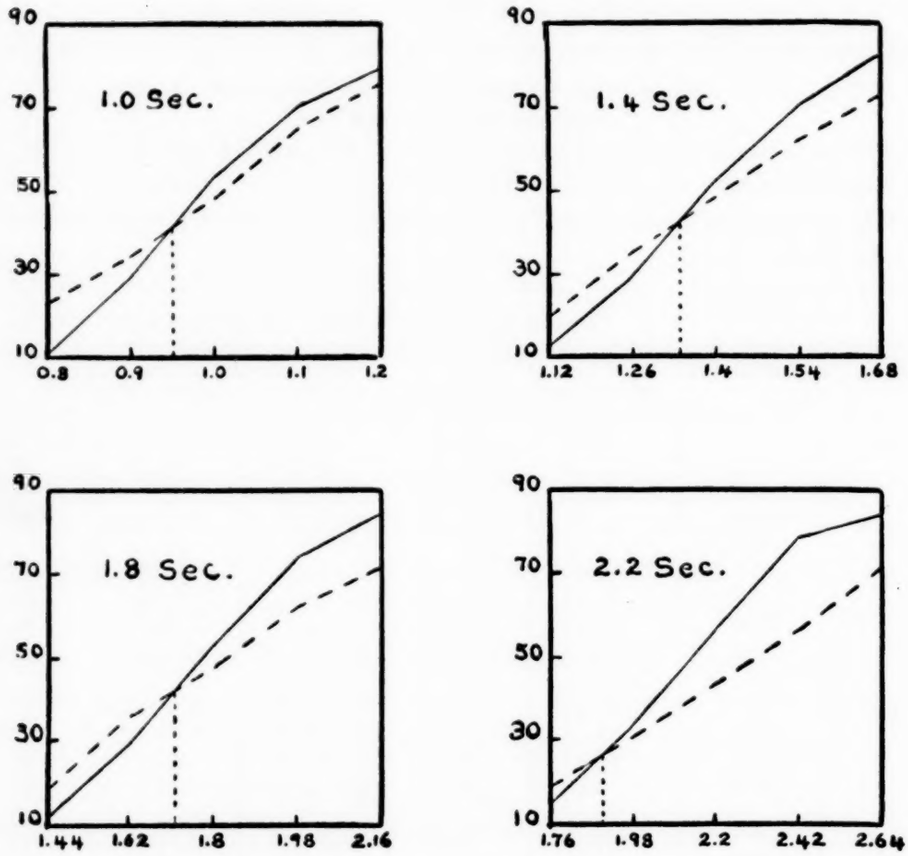

FIG. 3. Graphs illustrating the differences in percentages of judgments which signified 'variable longer' between the two orders of presentation, data from Table II. Solid lines show the standard first (S-V) order, broken lines show the variable first (V-S) order. Each percentage is based upon 780 judgments, 30 from each of 26 subjects. Same subjects served throughout. Variables ranged from 20 percent shorter to 20 percent longer than the standard.

instead of the hoped for verification, of the previous results was a baffling one. The causes, however, must have been in some way connected with some of the differences in conditions between the two experiments. Most of these differences were so small, or were of such a nature that little consideration need be given them. Others, it appeared, probably had important connections with the problem. They may be listed as follows:

1. Differences in apparatus.

2. Five of the fifteen subjects in the first experiment served individually while all subjects of Experiment II sat in groups.

3. Difference in the number of subjects.
4. Order in which the standards were used.
5. Total range of durations involved.
6. Difference in the number of trials per sitting.
7. Difference in the range and distribution of variable with respect to the standard.

1. During the first experiment the electrical sound making set-up was substituted for the mechanical gate apparatus. However, the pitch, loudness and quality of the tones as heard by the subject were practically the same in the two cases. Moreover, the results from the original five subjects were in no consistent way different from those from the other ten with whom the electrical apparatus was used.

2. The fact that the first five subjects of Experiment I served individually also seems not to be a cause for our contradictions for the reason given above, *i.e.*, that the results from them were not essentially different from those of the remaining ten who sat in groups of four and six respectively.

3. Even though the number of subjects in either case, and especially in Experiment I, is relatively small in view of the rather wide variations between individuals with respect to their time-order errors, it is difficult to see any possible connection between the number of subjects used and the appearance of the four new indifference points in Experiment II.

4. In the first experiment the five individual subjects and the group of four, nine in all, were given the ascending series first followed by the descending series while the other group of six subjects was given the descending series first. Again this seemed not to be an important factor since no consistent difference was observable in results under the two conditions.

5. The difference in total range of durations covered in the two experiments is believed to have considerable bearing upon our present problem. The subjects of Experiment I began their series of sittings with durations either much shorter or much longer than the durations used in Experiment II. It follows that by the time those standards which were used in both experiments were reached, namely, standards

from 1.0 to 2.0 sec., the subjects used in Experiment I had had quite different previous experience than those used in Experiment II, since the former group had had hours of comparison of durations not used at all in the second experiment. While the exact manner of operation of this difference in previous experience is not clear, it nevertheless appears quite probable that it may have been an important factor in producing the discrepancy between the data of Experiment I and those of Experiment II. No attempt to obtain data which would unravel the discrepancy was made because it seemed better to pass on to an experiment where previous experience was controlled by abolishing it altogether, and where also the length of the sittings was shortened.

It should be noted here, however, that rather conclusive evidence was obtained that the results with any one standard are greatly affected by the previous experience of the subject with other standards. This evidence, presented as Experiment IV, consists in the fact that a large group of subjects who were tested with a standard of 0.85 sec., with the identical, asymmetrical series of variables used in Experiment I, but without any previous practice in the comparison of tonal durations,[4] gave radically different results than those obtained in Experiment I.

6. In both experiments large numbers of judgments were made in each sitting (100 in Experiment I and 150 in Experiment II). It is true that the difference between the two in that respect was quite large but it appears very doubtful that that difference, as such, could account for the discrepancy between them since in either case the number was sufficiently large to permit adaptation changes to take place.

While difference in length of sitting may have little to do with the difference between Experiments I and II, it nevertheless seemed quite possible that the experience of repeatedly judging the same standard-variable combinations during a

[4] They had been given by Professor Woodrow one sitting of 120 trials in the comparison of empty intervals with a standard of 1.0 sec. two days earlier, and the Seashore phonographic time and rhythm tests on the preceding day; but at any rate their previous experience in temporal comparisons was radically different and far less varied as regards range of standards employed than that of the subjects of Experiment I.

given experimental hour might produce a marked effect upon the constant errors. If this were true then the C.E.'s at the beginning of a long sitting would ordinarily be somewhat different than at the end.

In order to get some indication as to whether this hypothesis was correct, the judgments obtained in the first sitting of the twenty-eight subjects of Experiment II with the 1.0 sec. standard were retabulated. The first sixty judgments, thirty for each order of presentation, were compared with the last sixty judgments. The C.E.'s in terms of difference in percentages of 'variable longer' judgments between the two orders were as follows:

	0.8 sec.	0.9 sec.	1.0 sec.	1.1 sec.	1.2 sec.
First 60 judgments.....	0.0	−6.0	−5.2	−5.0	−5.0
Last 60 judgments.....	+5.0	+4.1	−3.8	−4.2	−1.0

These values show that in this particular instance a change took place in the character of the discriminating response as experience was gained in making the judgments. At the beginning of the sitting, there was a tendency in the case of all variables except the shortest (0.8 sec.), to overestimate the second (negative C.E.'s). At the end of the sitting the situation was somewhat different. The relationship between the two orders of presentation had become quite similar to that shown in the data for the whole sitting. An indifference point had appeared which was not present at the beginning, and this new indifference point was just below the length of the standard. It thus appears that long sittings may be responsible for the appearance, as in Experiment II, of an indifference duration slightly below the length of the standard, at least when the standard is not too much longer than the indifference point which existed at the beginning of the sitting.

Moreover, in view of the effect of practice which, as indicated in the above tabulation, is to raise the indifference point, it appeared probable that what may be termed the true indifference duration, that is, the indifference duration before the time-order errors have been influenced by pro-

longed practice, should be considerably shorter than any of the values therefor obtained in Experiments I or II. It even appeared possible that with more reliable data it might be found to be of the same length as for empty intervals. Blakely [5] found an indifference point for empty intervals of approximately 0.625 sec. using the same method as was used in the present study. More recently Woodrow [6] has reliably shown an indifference interval of that same length using the method of mean error. With these considerations in view a further attempt, described as Experiment III, to locate the indifference point for tone-filled intervals was made.

7. The last of the differences listed above between the two experiments is that in the range and distribution of the variables with respect to the standard. In the first experiment the five variables ranged from 0 percent to 20 percent *longer* than the standard. In the second experiment they ranged from 20 percent *shorter* to 20 percent *longer* than the standard. In the first case the total range of differences in a given set of durations was only 20 percent of the standard and the standard was the shortest length of the series. In the second case the total difference range was 40 percent of the standard, and the standard fell at the exact center of the range of lengths. While this difference in the range of variables used in the two experiments may very well have had an influence on the difference in results, it is not apparent how it would help to account for them.

The most probable conclusion, then, is that the peculiar finding of several indifference durations in Experiment II is largely a result of the effect of the long sittings; while the difference in results between I and II is to be regarded as mainly due to the effects of the radically different series of standards employed and the effect exerted on the trials with any one standard of the preceding work with other standards.

Experiment III

In this experiment the factor of experience was minimized by using in the case of each standard an entirely different

[5] Blakely, W. A., unpublished doctor's thesis, University of Illinois, 1933.

[6] Woodrow, H., The temporal indifference interval determined by the method of mean error. J. of Exper. Psychol., 17, No. 2, April 1934.

group of subjects and giving only short sittings with each standard. The standards studied and the number of subjects that served with each were as follows:

Standard	No. of Subjects
0.625 sec.	100
0.8 sec.	100
1.0 sec.	74[7]
1.6 sec.	86

The wider range of variables (20 percent shorter to 20 percent longer than the standard) was used in every instance. Only sixty comparisons, six for each standard-variable combination, were made with any one standard, except the 1.6 sec. standard, with which in the case of forty-five of the eighty-six subjects 150 comparisons were made. In all other respects exactly the same experimental conditions were maintained as in Experiment II.

Results.—The results from this experiment are summarized in Table III. The percentages of judgments signifying 'variable longer' for each order of presentation are shown together with the percentage differences between the orders for each variable (V-S minus S-V). These last mentioned values represent the time-order errors. The P.E.'s of these percentage differences and their critical ratios are also included in the table.

With the 0.625 sec. standard, the C.E.'s are all reliably positive indicating a tendency at these durations to underestimate the second in relation to the first. With the 0.8 sec. standard, the C.E.'s are again positive in every case. However, the difference for the longest variable, 0.96 sec., is so small as to amount practically to zero. This indicates the presence of an indifference point in the neighborhood of 0.96 sec. The 0.8 and 0.9 sec. variables with the 1.0 sec. standard gave positive differences while the 1.0, 1.1 and 1.2 sec. variables gave reliably negative differences. This change in the sign of the C.E. quite definitely indicates the existence

[7] Since the 1.0 sec. standard was used in the very first sitting of the 26 subjects in Experiment II the data of the first 60 trials from that sitting are combined with those from the 74 new subjects in this sitting, bringing the total number up to 100.

TABLE III

EFFECT OF ORDER OF PRESENTATION UPON PERCENTAGES OF JUDGMENTS INDICATING THE VARIABLE AS LONGER THAN THE STANDARD IN THE RESULTS OF EXPERIMENT III

The percentage differences between the two orders (V-S minus S-V) represent the time-order errors.

A different group of subjects made only one short sitting with each standard.

Standard (Sec.)	Order of Presentation	Percentages 'Variable Longer' Judgments for All Variables				
		−20 percent	−10 percent	0 percent	+10 percent	+20 percent
0.625		(0.5)	(0.5625)	(0.625)	(0.687)	(0.75)
	V-S	19.33	34.83	54.33	71.92	81.17
	S-V	10.67	26.17	45.67	60.58	74.08
	Diffs.	+8.66	+8.66	+8.66	+11.34	+7.09
	$P.E._{D.}$	1.38	1.78	1.94	1.24	1.62
	$D./P.E._{D.}$	6.27	4.86	4.46	9.14	4.38
0.8		(0.64)	(0.72)	(0.8)	(0.880)	(0.96)
	V-S	18.65	35.62	52.98	69.84	81.55
	S-V	8.23	24.50	47.02	64.89	81.35
	Diffs.	+10.42	+11.12	+ 5.96	+ 4.95	+ 0.20
	$P.E._{D.}$	1.43	1.93	2.12	1.99	1.65
	$D./P.E._{D.}$	7.29	5.76	2.81	2.49	0.12
1.0		(0.8)	(0.9)	(1.0)	(1.1)	(1.2)
	V-S	13.83	28.42	46.04	69.75	80.50
	S-V	9.08	27.16	53.96	74.92	87.00
	Diffs.	+ 4.75	+ 1.26	− 7.92	− 5.17	− 6.50
	$P.E._{D.}$	1.24	1.74	1.94	1.74	1.43
	$D./P.E._{D.}$	3.83	0.72	4.08	2.97	4.55
1.6 (30 judgments from each subject)		(1.28)	(1.44)	(1.6)	(1.76)	(1.92)
	V-S	10.00	26.90	39.70	61.80	67.40
	S-V	19.00	35.40	60.30	81.00	90.30
	Diffs.	− 9.00	− 8.50	−20.60	−19.20	−22.90
	$P.E._{D.}$	2.07	2.74	2.91	2.63	2.33
	$D./P.E._{D.}$	4.35	3.10	7.08	7.30	9.83

of a time-order indifference duration for tone between 0.9 and 1.0 sec. Its most probable length is 0.92 sec. All differences with the 1.6 sec. standard are reliably negative. These facts may be clearly seen in Fig. 4. Naive subjects, then, in making comparative judgments of tonal durations, under the conditions here described, tend to underestimate the length of the second sound as compared with the first when they are 0.9 sec. or shorter in duration and to overestimate the second when they are 0.96 sec. or longer in duration. On the average, at a length of about 0.92 sec. the effect of secondness or firstness upon the comparative judgment is zero.

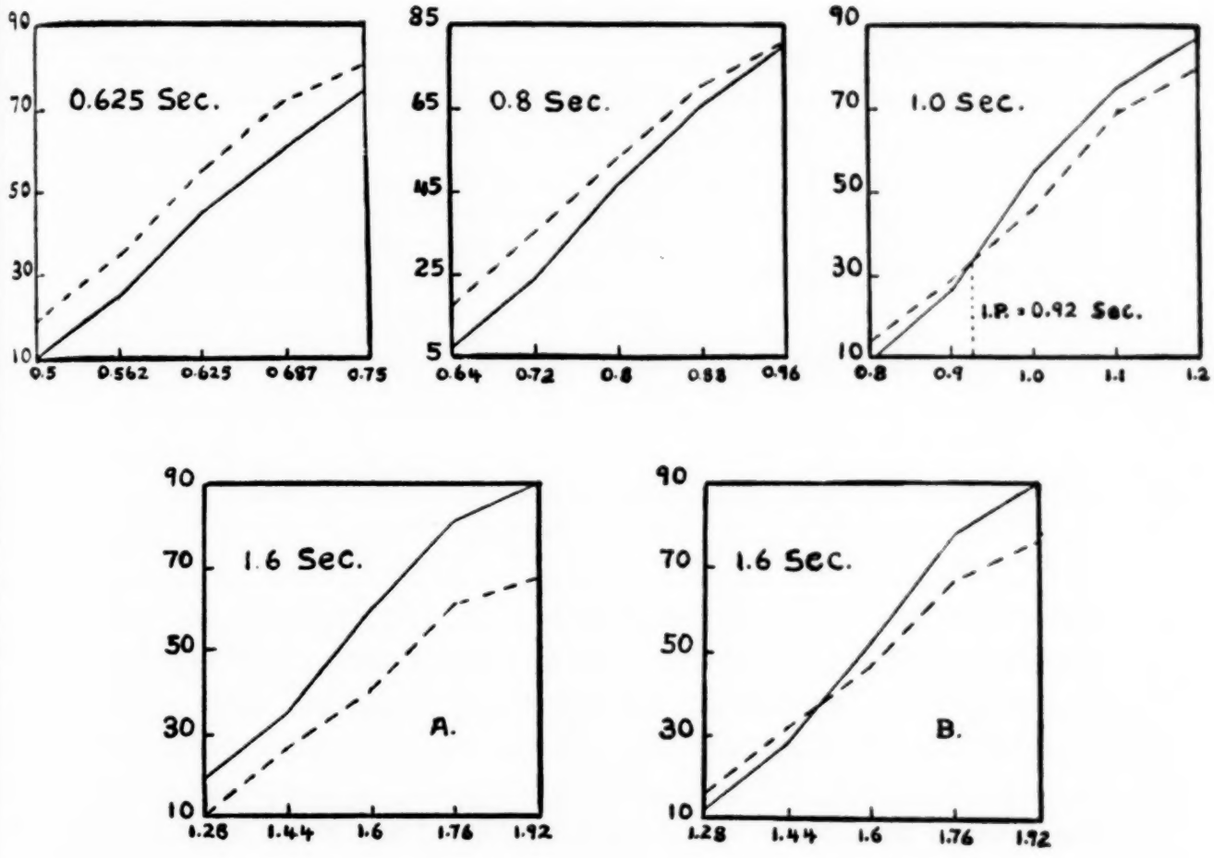

FIG. 4. Graphs illustrating the differences in percentages of judgments which signified 'variable longer' between the two orders of presentation. Solid lines show the S-V order and the broken lines show the V-S order. Different groups of naive subjects served in short sittings with each standard. *A* and *B*, represent the data presented in Table IV. *A* graphs are based upon only the first three judgments of each combination with the 1.6 sec. standard. *B* graphs are based upon all 15 judgments of each combination from 45 subjects.

It was mentioned above that forty-five of the eighty-six subjects who served with the 1.6 sec. standard were given a complete set of 150 presentations. In order to show more clearly the marked effect of experience obtained during the earlier trials of a given long sitting upon the average constant errors of that sitting all the results obtained with the 1.6 sec. standard are presented in Table IV. In the top part of the table are shown the results of the first thirty judgments only (three for each standard-variable combination) from all eighty-six subjects. At the bottom of the table the percentages are based upon the total results from the complete long

TABLE IV

SHOWING THE EFFECTS OF PRACTICE DURING A LONG SITTING UPON THE CONSTANT ERRORS WHEN ALL THE COMPARED DURATIONS ARE GREATER THAN THE INDIFFERENCE DURATION

Length of Standard, 1.6 sec.

No. of Judgments Each Subject	Order of Presentation	Percentages 'Variable Longer' Judgments for Each Variable				
		1.28	1.44	1.6	1.76	1.92
30 (86 subjects)	V-S	10.0	26.9	38.7	61.8	67.4
	S-V	19.0	35.4	60.3	81.0	90.3
	Diffs.	−9.0	−8.5	−20.6	−19.2	−22.9
150 (45 subjects)	V-S	15.9	31.8	47.8	67.1	75.7
	S-V	12.7	28.8	52.2	78.2	89.5
	Diffs.	+3.2	+3.0	−4.4	−11.1	−13.8

sitting of the forty-five subjects. In the former case all C.E.'s are reliably negative. In the latter case the situation is, in general, strikingly similar to that found with each standard in Experiment II. Both positive and negative C.E.'s are shown and an 'indifference point' appears between the second and third variable lengths. (See Figs. 3 and 4, *A* and *B*.) In the light of these results the conclusion seems warranted that the peculiar outcome of Experiment II (different indifference point for each standard) was very largely a result of the effect upon the total results of each sitting of repeatedly judging the same standard-variable combinations.

The matter which remains to be considered is the discrepancy between the results of Experiment I and those of Experiment III. In the former case an indifference duration in the neighborhood of 1.67 sec. was obtained. In the latter it was found to be 0.92 sec. The only differences between the two experiments which seemed at all likely to be connected with the discrepancy were those in (1) the number of subjects used, (2) the range and distribution of variables in relation to the standard, (3) the number of judgments per sitting, and (4) the number of standards with which the subjects had experience.

The difference in the numbers of subjects used in the two experiments was large. As to just how important it was in

producing the differences in results cannot be determined with data of the present study. Individual differences in time-order errors are undoubtedly large. It may possibly be that the relatively small number of subjects used in Experiment I was partly responsible for the great length of that indifference point. However, the fact that none of the earlier investigators, although disagreeing widely as to the length of the indifference point for tone filled intervals, found it to be less than 1.5 sec. may be significant in this connection. Thorkelson,[8] who was one of the first to study tonal length discrimination, used sounds from 1.5 to 15.0 sec. in length. He apparently got throughout predominantly negative C.E.'s. Wrinch,[9] on the other hand, found only positive C.E.'s with sounds from 0.25 to 1.5 sec. in duration. The results of these two investigators, when considered together, suggest an indifference length in the neighborhood of 1.5 sec. Quasebarth [10] got positive C.E.'s at 2.0 sec. and negative ones at 4.0 sec., suggesting an indifference point between those two lengths. Hulser [11] found one at about 1.8 sec. It does not seem likely that all these cases of long indifference points were due to the chance selection of subjects whose individual indifference points were exceptionally long. Some other factor or factors, it appears, which may have been operative in the work of the earlier investigators, were more probably responsible for the long indifference point in the results of our first experiment. It is believed that some light is thrown upon the problem by the following test, Experiment IV.

Experiment IV

Two new groups of naive subjects, 58 men and 55 women, university students, were given one sitting each with a standard length of 0.85 sec. The variables ranged from 0 percent to 20 percent longer than the standard—the same narrower range that was used in Experiment I. As in all the work of

[8] Thorkelson's work was reviewed by Meumann, *Philos. Stud.*, 1893, **8**, 432–441.

[9] *Philos. Stud.*, 1903, **18**, 274–327.

[10] *Arch. f. d. ges. Psychol.*, 1924, **49**, 379–432.

[11] *Arch. f. d. ges. Psychol.*, 1924, **49**, 363–378.

this study the standard was given second in one half of the presentations. The combined results from the 100 presentations given each of the two groups are summarized in Table V.

TABLE V

EFFECTS OF ORDER OF PRESENTATION UPON PERCENTAGES OF JUDGMENTS INDICATING THE VARIABLE AS LONGER THAN THE STANDARD AS SHOWN IN INITIAL TRIALS ONLY AND IN THE COMPLETE SITTING

The percentage differences between the two orders (V-S minus S-V) represent the time-order errors.

Length of standard, 0.85 sec.

Total number of subjects, 113.

No. of Judgments Each Subject	Order of Presentation	Percentages 'Variable Longer' Judgments for Each Variable				
		0.85	0.8952	0.935	0.9775	1.02
10	V-S	53.5	70.3	61.1	65.9	85.8
	S-V	46.5	66.8	66.8	68.1	88.5
	Diffs.	+7.0	+3.5	−5.7	−2.2	−2.7
100	V-S	54.6	62.5	69.1	73.6	80.2
	S-V	45.4	61.7	70.5	80.8	86.8
	Diffs.	+9.2	+0.8	−1.4	−7.2	−6.6

These results are to be compared with those from Experiment I, Table I. There the C.E.'s are all positive for all durations less than 1.65 sec. In Table V a change in the sign of the C.E. is indicated between 0.89 and 0.935 sec., which again verifies the existence of an indifference point for tonal durations of approximately 0.92 sec. in naive, unpracticed subjects.

Since the same set of variables, consisting only of durations longer than the standard, was used in the present test as in Experiment I, and since in the present test the same indifference duration was obtained as in Experiment III, it seems rather clear that the discrepancy between the results of Experiments I and III could not be due to the difference in the range of variables used in those two experiments. The most plausible explanation, therefore, of the discrepancy between the results of Experiments I and III is the fact that in the first experiment the subjects when tested near the indifference point had already had long hours of practice at

discriminating durations of other lengths, while in Experiment III no subject was given more than one short sitting with just one standard. The evidence at hand, then, seems to indicate that differences in experience, due to differences in the range and order of standards used, and to differences in the length of sittings, are probably important causes of the great amount of disagreement among investigators of the past with respect to the time-order indifference point for tonal durations.

That in the present test there does not appear the change with practice in the time-order errors which was noted in the two previous cases (1.0 sec. standard, Experiment II, and 1.6 sec. standard, Experiment III) is in all probability due to the fact that in the present test the indifference duration, even when calculated from the initial trials, lay near the middle of the range of variables. This is shown by the fact that the two shortest variables initially gave positive time-order errors while the three longest gave negative time-order errors. In the two cases in which a shift in the time-order errors was observed, on the other hand, the initial trials showed positive time-order errors in the case of none of the variables, so that the indifference point lay either below the shortest variable or (in the case of the 1.0 sec. standard) coincided with the shortest variable.

Summary and Conclusions

An experimental study was made of the time-order errors in the comparison of short tonal durations with the purpose of determining the duration of tone at which these errors on the average become zero. This time-order indifference duration has been a matter of wide disagreement among the experimenters of the past.

Great care was exercised and much time was consumed in the development of an apparatus by means of which continuous sounds free from clicks and noises could be accurately presented and precisely described as to duration.

University students and staff members of the Department of Psychology, University of Illinois, 524 in number, served

for varying lengths of time as subjects. From these subjects some 99,480 judgments were obtained upon durations ranging from 0.2 to 36.0 sec. From these data as summarized in the foregoing tables the following conclusions were drawn:

1. With naive, unpracticed subjects the time-order indifference point for tonal durations is approximately 0.92 sec. A relative underestimation of the second member of each pair for all durations of 0.90 sec. and shorter, and a relative overestimation of the second for all durations of 0.96 sec. and longer were definitely shown. This indifference point is much shorter than those obtained by earlier investigators using tone filled intervals.

2. Experience obtained by subjects during a long series of experimental sittings in which a wide range of standards is covered apparently affects the time-order errors in such a way as to cause a marked change in the obtained length of the indifference interval. The average results from fifteen subjects who served throughout a lengthy series of sittings showed an indifference duration between 1.6 and 2.0 sec., or 0.7 to 1.1 sec. longer than that for naive subjects.

3. With series in which the five variables range from 20 percent shorter to 20 percent longer than the standard the C.E.'s undergo a marked change during a given long sitting as many repeated judgments of the same standard-variable combinations are made. In cases where the durations are all greater than the indifference point (0.92 sec.), and consequently where the initial C.E.'s are all negative, this change is such that the total results of the sitting show positive C.E.'s in connection with the shorter variables. An 'indifference point' is thus produced where none existed in the early stages of the sitting. It was therefore necessary, in order to minimize these effects of experience, to use large numbers of subjects with each standard and to consider only the early judgments of the sitting. Only in that way was it possible to obtain consistent and understandable results regarding the time-order errors and the indifference point for tone filled intervals.

(Manuscript received August 22, 1934)

INTERPOLATION EFFECTS WITH DIFFERENT TIME INTERVALS

BY J. GARTON NEEDHAM[1]

Directed errors appear among the judgments of nearly all types of stimuli presented for estimation or comparison. Numerous experimental studies have been concerned with the general fact that a certain class of these errors, the time-errors, varies as there is a change in the time interval separating two compared stimuli. In the attempt to ascertain the principles underlying this phenomenon (known as the p-function) various special experimental conditions have been devised. Prominent among these are the investigations of the 'pre-disposing' or 'contrast' effects which are to be noted in the displacement of the time-errors when extraneous stimulus presentations condition the judgment of the comparison stimuli.

Data obtained with the interpolation of a constant stimulus background, and with the interspersion of various intensive levels of stimulation, show that under such influences the judgment of the second of two stimuli with respect to the first tends to be directed in opposition to this extraneous influence. For example, with the presentation for comparison of two stimuli upon a continuous background more intense than the stimuli themselves, there appear a greater number of judgments of the second as 'Less' than would be expected otherwise; again, the effect of the presentation and judgment of a series of considerably more intense stimuli upon the subsequent judgment of a series of less intense stimuli is an increased relative number of judgments 'Less.' Moreover, further investigation shows that if under such experimental conditions as these the time interval between the comparison stimuli is increased, there results an enhancement of these

[1] National Research Council Fellow.

'pre-disposing' influences of the extraneous stimulus presentations.[2]

There is one body of experimental data which should, upon logical grounds, correspond to the above-mentioned findings. If there is inserted between the two stimulus presentations which comprise a comparison pair an extraneous stimulus of an intensive value considerably greater or less than that of the stimuli themselves, it would be consistent with the above discussion for the judgment of the second stimulus to be displaced in a direction opposite to that of the interpolated stimulus. Further, it would be predicted that this interpolation-effect would become greater as the time interval between the comparison stimuli was increased. Such consequences would be revealed in a relatively greater positive time-error (more relative preponderance of judgments 'Less') with the interpolation of a more intense stimulus, and a relatively greater negative time-error (greater relative overestimation of the second stimulus) when the interpolated stimulus was less intense than the comparison stimuli. With a longer time interval this difference would become more marked.

In a lifted weight experiment by Guilford and Park,[3] in which a heavier or a lighter weight was interpolated between the two comparison weights, measures comparable to those of the time-error indicate the presence of these opposite effects, and therefore the results are consistent with other information concerning the time-error phenomena. On the other hand, although Pratt's results confirm this opposition of the two interpolation effects, the data indicate also that the errors in the comparison judgment under normal conditions (with no interpolated stimulus) are inconsistent if it is to be assumed

[2] The present report is supplementary to a preceding paper, The effect of the time interval upon the time-error at different intensive levels, this JOURNAL, **18**, 1935, 530–543. The majority of the bibliography involved here is cited there; the original p-function discussion is that of Köhler; Lauenstein is responsible for the data upon judgments with a continuous background; 'contrast' effects have been studied by Woodrow, Pratt, and Needham.

[3] J. P. Guilford and D. G. Park, The effect of interpolated weights upon comparative judgments, *Amer. Jour. Psychol.*, 1931, **43**, 589–599.

that the normal case is the extreme or limiting instance of interpolation of a weak or light stimulus; the time-error is less positive under normal conditions than with the interpolation of a stimulus of less intensity.[4]

Pratt's results probably involve a complication in rather than a contradiction of the general information concerning time-error phenomena. The purpose of the present paper is to indicate briefly the nature of further complications which derive from the experimental effects of interpolated stimuli. It was argued above that the opposition of interpolation effects should become greater with an increase in the time interval between the two comparison stimuli. Upon two occasions the writer has had the opportunity to obtain a limited amount of information concerning this matter. In both cases p-functions of the time-error were determined with the inclusion of interpolated stimuli either more or less intense than the comparison stimuli. There were certain gross differences in the experimental conditions in the two instances, and these differences are reflected in the results. However, the data are completely unequivocal as to the general direction of the displacement of the interpolation effects with a variation of the time interval; consequently they are presented at this time in order that their general characteristics may be available to discussion.

The results in question were obtained in experiments which were supplementary to investigations reported elsewhere, one (I) from the Harvard Psychological Laboratory,[5] the other (II) from the Princeton Psychological Laboratory;[6] the general conditions, apparatus, and so forth, are described in the references cited. In both cases the stimuli were auditory intensities. Other special features were:

[4] C. C. Pratt, The time-error in comparison judgments, *Amer. Jour. Psychol.*, 1933, **45**, 292–297.

[5] *Amer. Jour. Psych.*, 1934, **46**, 558–567.

[6] Jour. Exper. Psych., *loc. cit.*

	I	II
Observers:	Wo. Mr. A. Wolbarsht Mc. Mr. D. M. McGregor	We. Dr. C. H. Wedell Co. Mr. J. T. Cowles
Stimuli:	Clicks, loudness level not measured. Range of variation less than in II	Comparison stimuli 40 db. above RL; interpolated stimuli 20 db. greater or less
Stimulus duration:	Momentary	½ second each
Interval between comparison stimuli:	2, 4, 8 seconds	1, 3, 6 seconds
Interval between interpolated and each comparison stimulus:	1, 2, 4 seconds (approximately)	¼, 1¼, 2¾ seconds
Total number of judgments:	900	600

The results are presented graphically in Fig. 1. A positive value of the time-error indicates a preponderance of judgments of the second comparison stimulus as 'Weaker,' and a negative value means that the 'Louder' judgments predominate. As concerns the respective levels of the values for the two interpolation conditions, the graphs of Fig. 1 agree with previous findings; with a loud interpolated stimulus the time-error is either more positive or less negative, while with a weak interpolated stimulus the time-error is either more negative or less positive. The fact that the curves tend to lie at different levels in the two experiments is probably attributable to some undetermined factor in the difference between the two sets of experimental conditions.

The crucial significance of the p-function curves of Fig. 1 is the fact that with a lengthened time interval there is a decrease in the opposition between the two interpolation-effects. Rather than receding, the two curves draw closer together for each of the four observers; the time-error discrepancies become less marked. With the loud interpolated stimulus there are fewer rather than more 'Weaker' judgments at the longer intervals; when the weak stimulus is interpolated, the effect of a longer interval is a relatively smaller number of judgments 'Louder.'

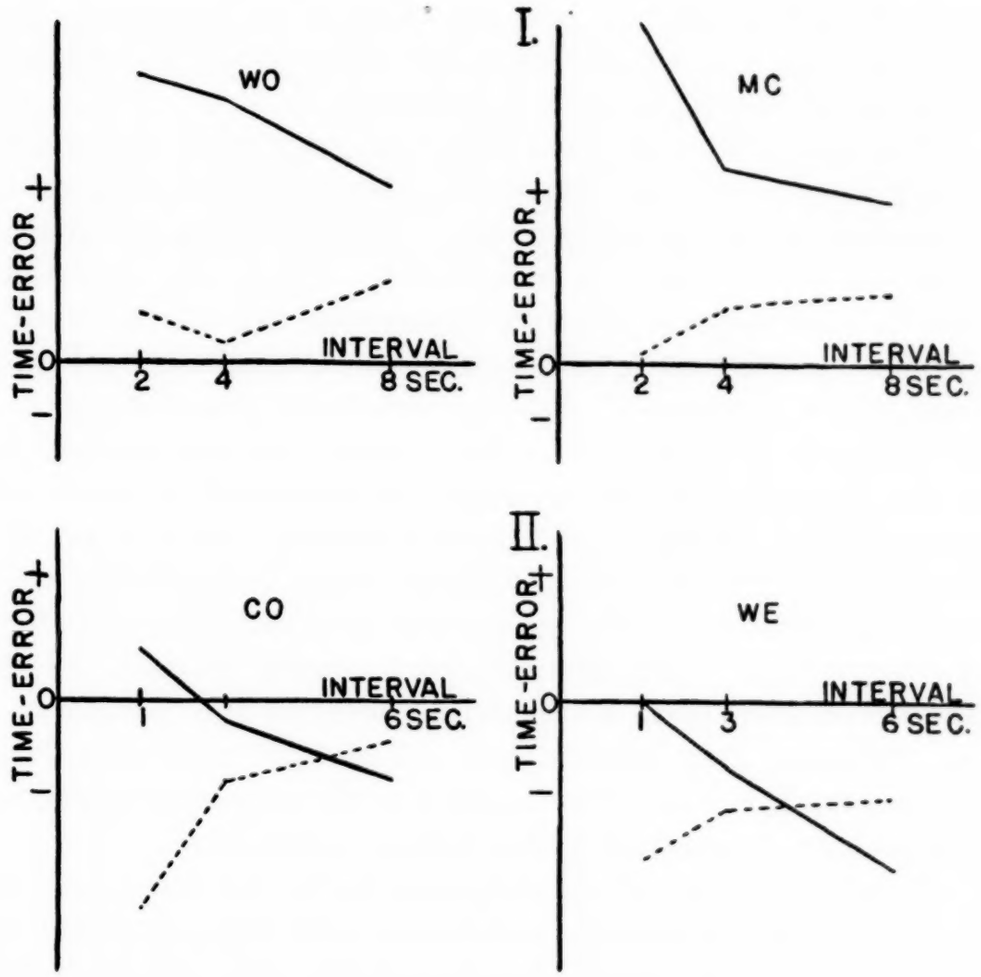

FIG. 1. p-functions under two conditions of interpolation.
—— Interpolated stimulus—Loud,
---- Interpolated stimulus—Weak.

The p-function curves for both observers in II intersect, while those for the observers in I do not, and there are other differences between the data for the two experiments, such as the levels at which the curves lie. However, so numerous were the differences in the experimental conditions that such discrepancies are to be expected. The nature of the stimuli was not the same in the two cases. The temporal dimensions upon which the p-functions were determined were different. The duration of the interpolated stimulus was greater in the case of II, and its effect might be expected to be altered by this condition. These are matters for further investigation. Meanwhile, despite the discordant features of the data, there

is no mistaking the fact that the effect of an increase in the time interval is a diminution of the effect of an interpolated stimulus upon the comparison judgment.

The importance of this result with respect to theoretical considerations of the time-error phenomena in general is evaluated in the preceding paper. There remains one generalization which may be mentioned. There are available certain sophisticated theories concerning the effects of the various 'pre-disposing' influences as determinants of the comparison judgment. Among these the principle of assimilation is prominent. In the present case, the opposition of the interpolation effects might be explained in terms of assimilation. However, if the determining factor is an assimilative one, then there should be 'more' assimilation with the longer time intervals; the present data directly contradict this hypothesis. The interpolation condition seems to be an instance wherein with a longer time interval the 'pre-disposing' influence may become less rather than more effective, and as such it is a signal for caution in the general application of principles determined under limited conditions.

A naïve analysis of the judgment under the interpolation conditions can be entirely consistent with the experimental findings. As the time-error lengthens, the interpolated stimulus obviously becomes more and more remote (temporally and phenomenally) from the two stimuli upon which the judgment is based, and its effectiveness as a 'disturbing factor' certainly decreases. The purpose in suggesting such an analysis as this is not a denial of such concepts as that of assimilation; rather, it is intended to suggest the need of a thorough analysis of the process of judgment in more terms than those of easily adapted physiological postulates. The comparison process, together with other forms of the so-called higher mental processes, is not a simple affair; the determinants are numerous, variability is the rule rather than the exception, and a limited physiological hypothesis, such as that of a 'sinking trace' (Köhler) or of assimilation (Lauenstein), implies a constancy and a simplex determinism which belie the true state of affairs. Indeed, even the 'pre-dis-

posing' influences which are discussed above are relative effects at best; their presence is noted in the appearance of *more* of a certain class of judgments, never in a totality of that class of judgments. The qualification of 'one-way' hypotheses is demanded still further by such experimental results as those obtained under conditions of interpolation.

(Manuscript received August 7, 1934)

BILATERAL TRANSFER OF THE CONDITIONED KNEE-JERK

BY JAMES J. GIBSON AND LUCILE HUDSON

Smith College

The experiment to be described was performed in order to follow up an earlier demonstration which showed that the conditioned withdrawal response of one hand to an electric shock transferred to the other hand, even though the second hand had not been given training.[1] We wished to find out whether a conditioned knee-jerk in one leg is accompanied by the formation of a similar but latent conditioned response in the other. If such a latent response could be elicited under appropriate circumstances, additional evidence would be provided for the proposition that conditioned responses of the skeletal musculature are related to skills and habits characteristically showing transfer.

The method used to condition the knee-jerk in the trained leg was as follows. The subject was instructed to put his entire attention upon reacting as quickly as possible to the stimulus of a combined light and buzzer by pressing a grip-key which he held in the hand on the side to be trained. A ready signal was given, the stimulus and the reaction followed; .28 sec. after the beginning of the stimulus, a blow was delivered to the patellar tendon. In short, the component events of a voluntary reaction, serving incidentally to facilitate the reflex, were employed as the conditioned 'stimulus.' In the event of a very long reaction-time, the hand response would occur slightly after the blow on the tendon, but normally it would occur before.

This method was suggested by Dr. Alexander Mintz who uses it informally, without apparatus and giving signals verbally to demonstrate conditioned knee-jerks. Many sub-

[1] James J. Gibson, E. G. Jack, and G. Raffel, Bilateral transfer of the conditioned response in the human subject, J. Exp. Psychol., 1932, 15, 416–421.

jects show conditioning after 4 or 5 trials. A very similar procedure was used by Schlosberg [2] who found that when the knee-jerk was facilitated by a concurrent voluntary reaction, conditioning occurred much more readily. Schlosberg, however, varied the kind of voluntary reaction required of his subject on successive trials, and used a shorter interval between his conditioned stimulus and the blow on the tendon. He did not believe that the voluntary response acted as a part of the conditioned stimulus. Our method is also related to that of Hudgins [3] who conditioned the pupillary reflex to a voluntary clenching of the hand, and to that of Stephens,[4] who conditioned the withdrawal-response of one hand to a voluntary reaction of the other hand, finding that either stimulus alone would elicit the other response in addition to its own, the voluntary reaction being the more 'dominant' of the two arcs.

Apparatus.

The subject sat in an adjustable seat with padded leg-rests. The legs could swing freely. A screen prevented any view of the legs or the apparatus. Two identical pendulum-type reflex hammers, released by small electromagnets and adjusted to fall through 45 degrees, hung in front of the subject's knees. After preliminary tests, only the hammer for the trained knee was used. The experimenter caught and reset the hammer after each blow. No method of recording the knee-jerks graphically was employed; the kicks obtained were easily observable and the experimenter had only to note the occurrence or non-occurrence of conditioned responses. A background of white cardboard ruled with cross-lines under the subject's feet made a very slight movement perceptible. The reflex hammer was released by pressing a key which simultaneously broke the electromagnet-circuit and made the circuit activating a 6-volt light bulb and a buzzer, in parallel, which were attached to the screen in front of the subject. The subject reacted by clenching a wooden grip-key, breaking the latter circuit and shutting off the light and the buzzer. About 10 lb. of force had to be exerted in order to compress the key.

Procedure.

At the first sitting, after the subject had been seated in the chair, he was given the grip-key and allowed practise with both hands individually in clenching it. The

---

[2] H. Schlosberg, A study of the conditioned patellar reflex, J. Exp. Psychol., 1928, 11, 475–6, 481.

[3] C. V. Hudgins, Conditioning and voluntary control of the pupillary light reflex, *Jour. Gen. Psychol.*, 1933, 8, 11 f.

[4] J. M. Stephens, The conditioned reflex as the explanation of habit formation: II. The operation of a higher-order reaction and a lower-order reaction in close succession, *Jour. Gen. Psychol.*, 1934, 10, 219–227.

experimenter demonstrated the light-buzzer stimulus and then told the subject that his task was to react to it as quickly as possible by gripping and briefly holding down the key, thereby cutting short the light and buzzer. A ready signal would always be given. The subject was given training with both hands in this reaction. His legs were carefully watched for a response to any of these stimuli. None appeared in any subject. Next the stimulus-hammers were adjusted for each leg. Several knee-jerks were elicited and the other leg was observed for any response such as a crossed reflex. On one occasion, a crossed response was noted. It appeared in a subject who subsequently showed no conditioning. The subject was then given the grip-key to hold in the hand on the side selected for training, the instructions were repeated, and the training-series was begun. If the subject was naive, he was told that the purpose of the experiment was to observe the quickness of his reaction when accompanied by a knee-jerk.

The first training-series consisted of 5 combined hand-reactions and reflex knee-jerks, after which a test for conditioning was made. The experimenter caught the stimulus-hammer before it struck and noted any response of the leg. Subsequent training-series consisted of 15 combined stimulations, succeeded by a test, until stable conditioning appeared, after which reinforcements and tests were given irregularly. If no evidence of stable conditioning appeared in the first sitting, the subject was dismissed. Stimulations were given at a rate of about 6 per minute. Frequent rest periods were introduced.

When two successive tests had yielded conditioned responses, the first test for a transferred conditioned response was usually made. The subject was told, "Now we shall work on the other leg," preparations were made as if to do so, *and the grip-key was transferred to the subject's other hand.* The subject then reacted to the stimulus by clenching the grip-key on the transfer-side, no blow was delivered to the tendon, and any movement of the leg was noted. At no time during the experiment were combined stimulations given on the transfer-side. When transferred conditioned responses had been elicited several times, an attempt was made to obtain in succession a conditioned response, a transferred conditioned response, and another conditioned response. Training was continued with the goal of obtaining eventually 15 conditioned responses in succession, the middle 5 from the transfer-leg. Sixteen subjects were used, instructors and students, of whom 7 were naive.

Results.—The results are given in Table I which shows the outcome of individual tests for conditioning and transfer up to the 65th double stimulation, which was usually reached at the end of the first sitting. The results of subsequent tests are given in the last column of the table in percentages of successful conditioned and transferred responses. It should be noted that 10 subjects did not show signs of stable conditioning and were dismissed after the first sitting.

Of the 16 subjects, the first 12 listed in the table gave at least one conditioned response by the end of the 65th double stimulation. Nine showed conditioning after the 5th double stimulation. Five subjects gave at least one transferred

TABLE I

RESULTS OF TESTS FOR CONDITIONED RESPONSES AND TRANSFERRED CONDITIONED RESPONSES IN 16 SUBJECTS

Subject	Test 1 (5 stim.)	Test 2 (20 stim.)	Test 3 (35 stim.)	Test 4 (50 stim.)	Test 5 (65 stim.)	Succeeding Tests
E.G.....	C	C-T	C	C-T	C	C—100%; T—100%
E.S.*....	C	C-T	C-T	C-T	C	C—100%; T—100%
M.R.....	C	C-X	C-X	C	C-T	C— 91%; T— 85%
H.R.....	X	C	C-T	C	C	C— 96%; T—100%
M.H.....	C	X	X	X	X-X	C— 44%; T— 17%
D.B.....	C	C-T	C	C	X	C— 14%; T— 0%
M.S.....	C	X	C-X	X	X	Subject dismissed
B.....	C	X	C-X	X	X-X	Subject dismissed
K.M.*...	C	X	X	X	X	Subject dismissed
L.R.*....	C	X	X	X	X	Subject dismissed
S.*.....	X	C-X	X	X	X	Subject dismissed
F.B.....	X	C	X	X	X	Subject dismissed
J.S.....	X	X	X	X	X	Subject dismissed
W.*.....	X	X	X	X	X	Subject dismissed
M.B.*....	X	X	X	X	X	Subject dismissed
M.P.*....	X	X	X	X	X	Subject dismissed

C—Conditioned response elicited.
T—Transferred conditioned response elicited.
X—Failure to obtain conditioned (or transferred) response.
*—Indicates naive subject.

conditioned response during the first sitting, and a sixth gave several scattered ones during subsequent training. It is the first four subjects, however, whose results are of greatest significance for our problem. They manifested stable conditioned responses from the beginning. In these subjects *not only were transferred conditioned responses obtained, but they occurred as consistently, on the whole, as the original conditioned responses themselves.* Three of these subjects were trained to the stage where they gave 15 conditioned responses in succession, the middle 5 of which were obtained from the transfer-leg. The fourth subject reached a level of 9 such successive responses, lack of time preventing further work. It may be concluded, then, that in those subjects in whom conditioning of the knee-jerk was strongly established, an equally stable conditioning of the other untrained leg had been set up, which showed itself when the experimenter said, "Now we shall work with the other leg," and when the other hand was used for the voluntary reaction.

The Result of Changing the Conditioned Stimulus-Complex. —The conditioned 'stimulus' in our experiment had to be stated broadly as the component events of a voluntary hand-reaction to a light and buzzer. It remained to be seen whether a single component of this complex, *e.g.* the clench of the hand alone or the light-and-buzzer alone, was functioning as the adequate stimulus for the conditioned and the transferred response. If no single component was the active agent, then how much could the stimulus-complex be changed by alternately eliminating various components without making it ineffective? It was already known that the presence of the light-buzzer-stimulus was not essential to the conditioned response, for occasionally during training subjects had kicked before this stimulus appeared. It developed that they had given a premature hand-reaction immediately after the ready signal and the kick had followed.

In order to study this problem, each of the 4 subjects who showed stable conditioning was given an additional experimental sitting. After some reinforcement, tests were made for conditioned and transferred responses which were in every case found to be present. Then tests for the responses were made under the following modifications of the stimulus-conditions:

I. Hand-reaction eliminated. Subject listens passively to light and buzzer. Grip-key simply held in appropriate hand.

I*a*. Same, except that grip-key is not held.

II. Light-buzzer stimulus eliminated. S reacts to E's command, "Ready, clench!"

II*a*. Same except that set to react is changed. S reacts while herself saying aloud, "Ready, clench!"

II*b*. Same except that auditory stimulus is eliminated. S clenches while saying the words sub-vocally.

Tests were made 3 at a time for any one condition. Reinforcement was given after every 6 tests. The conditions were run through twice in a counterbalanced order; hence 6 tests were given each subject for each condition. The procedure was the same in testing the transfer-leg as in the trained leg.

The results are given in Table II which shows the number of times out of a possible 6 a subject kicked under each condition and with each leg. A total of 240 tests was made of which 159 were positive. It will be seen that different subjects yield somewhat different results, some being more

TABLE II

TESTS FOR CONDITIONING AND TRANSFER UNDER MODIFIED STIMULUS CONDITIONS

	I No clench. Key held		Ia No clench. Key not held		II No light-bz. Verb. command		IIa No light-bz. Self command		IIb No light-bz. Sub-vocal	
	Cond.	Trans.	Cond.	Trans.	Cond.	Trans.	Cond.	Trans.	Cond.	Trans.
H.R.....	4	0	5	0	4	2	0	0	0	0
E.G.....	2	6	1	6	6	6	6	6	6	6
M.R.....	2	0	2	0	6	6	6	6	5	6
E.S.....	6	6	6	6	6	6	6	6	0	6
Avg.....	3.5	3.0	3.5	3.0	5.5	5.0	4.5	4.5	2.8	4.5
Percent	54%		54%		88%		75%		60%	

affected by the changed conditions than others. In general, transferred conditioned responses were elicited under these conditions about as often as original conditioned responses. No essential difference appears between the two types, which fact indicates once more that the transferred responses are essentially similar to the original responses.

It is clearly evident that no single component of the stimulus-complex is responsible for the conditioned response. No component is common to all of the 5 conditions in the table, and every condition gave some responses. We may infer that the kick of the leg has been conditioned to the whole complex of stimulating events incidental to the voluntary hand-reaction. This conclusion is the more probable when one notes that every one of the 5 conditions representing some degree of change is less effective in producing responses than the original stimulus-complex. Some of the changed conditions, however, appear to be more effective than others. Although the averages and percentages at the bottom of the table should not be granted too much reliability in view of the individual differences, it is perhaps worth noting that Condi-

tion II, in which the only change was the substitution of the experimenter's command for the light and buzzer, yielded the highest proportion of responses. All of the II conditions gave higher percentages than the I conditions indicating probably that, insofar as they are separable, the clenching response was a more important factor in the total stimulating complex than the auditory-visual stimulus which released it.

Discussion.—Out of 16 subjects, 4 who showed stable conditioned knee-jerks in the trained leg also showed equally stable conditioned responses in the untrained leg. We must suppose that the learning which had occurred was effective for either leg—that there was a bilateral transfer of such learning. Gibson, Jack, and Raffel [5] demonstrated that when there had been established a conditioned withdrawal response of the middle finger of one hand, using an electric shock as the unconditioned stimulus, a similar conditioned response could usually be elicited in the middle finger of the other hand when it was placed on the electrode and when the conditioned stimulus, a buzzer, was sounded. An explanation of the transfer was suggested in terms of a generalized system of avoidance responses, effective for various parts of the body, which as a unit might have become conditioned to the buzzer stimulus. What the subject had learned, we supposed, was an automatic habit of avoiding shocks in that particular situation.

A hypothesis of this type is not so easily applicable to the facts of the present experiment. A generalized habit of kicking is not as plausible as a generalized habit of withdrawing from an electrode because of the absence in the former of any apparent unifying motivational factor. Only by supposing that the knee-jerk is an artifically produced manifestation of what is in reality a unit in a bilateral system of postural responses, could one meet this difficulty. Crudely, if the knee-jerk is a by-product of the standing-up response to an incipient falling-down stimulus, it is not hard to see how conditioning of this response in one leg might transfer to the other.

[5] J. J. Gibson, E. G. Jack, and G. Raffel, *op. cit.*

Irrespective of what theory is proposed to explain the transfer of the conditioned knee-jerk, another equally puzzling problem remains to be faced. If the conditioning which has occurred is effective in both legs, why is it that both legs do not kick at the same time? Over 300 conditioned knee-jerks were observed in the course of this experiment and of these only 8 were accompanied by a simultaneous kick of the other leg, 6 being given by one subject (E.G.). There is evidently present a fairly efficient mechanism for releasing the previously latent response and inhibiting the previously overt response when the experimenter says, "Now we shall work with the other leg," or "Now we shall go back to the first leg." It will be remembered that when these shifts were made, the grip-key was transferred to the appropriate hand. The conditioned stimulus-complex was, accordingly, not identical for the transferred and the original response. This change in the stimulus-complex was probably an important factor making for the unilateral character of the conditioned responses. That it was not an essential factor, however, is proved by the obtaining of unilateral responses in either leg to the stimulus of the light and buzzer alone, the grip-key not even being held in the hand (Condition I*a*). Here the isolating factor could only have been the result of the verbal announcement made by the experimenter. The conclusion is inescapable that something in the nature of an anticipatory set was operating to release and inhibit reciprocally the responses of the two legs.

Despite this dependance on set, the conditioned knee-jerks cannot be said to be fully intended or willed, *i.e.* to owe their existence to a specific set to respond, as did the accompanying hand-reactions. The conditioned knee-jerks were at first not even expected or forseen. Subject H.R., for example, reported after the first transferred response, "I didn't expect that to happen at all, it just went." And later, "You feel completely dissociated, your knee just goes." At later stages of training the response might be anticipated but seemed to be no less automatic. Subject E.S. reported, "I know my leg is going to fly up when I clench my hand. Nothing could stop it."

Neither can the conditioned knee-jerks of this experiment be dismissed as merely due to suggestion, if by this is meant that subjects kicked because they understood the experimenter's unspoken preference, or because they knew that the experiment concerned conditioned responses. Several naive subjects showed conditioning before realizing that the absence of the blow on the tendon meant anything more than a mishap in the experimental procedure.

Whether the responses are or are not 'voluntary,' then, depends entirely on what meaning is assigned to that term. Introspectively they are not like voluntary reactions, or more strictly, are not like *unpracticed* voluntary reactions. By objective criteria, however, conditioned knee-jerks are indistinguishable from voluntary reactions in respect to latency and form,[6] and they are now seen to be linked with skills and manipulatory habits in showing bilateral transfer and in their capacity of being evoked separately, on either side of the body, by instructions from the experimenter. The precise relationship between conditioned responses and various types of voluntary acts cannot be determined without further experimental evidence. For one thing, a careful investigation is needed of the introspective similarities and differences between the knee-jerk, the conditioned knee-jerk, the voluntary kick, and the habit resulting from long practice of the voluntary kick, together with their susceptibility to voluntary inhibition. There is evidence that a *habitual* voluntary reaction is introspectively similar to a conditioned response,[7] and also that such a reaction may be 'conditioned,' *i.e.* elicited automatically, by an accompanying reflex response.[8] In the present experiment, as an after-thought, one subject (E.G.) was tested without warning for such cross-conditioning, that is, for the tendency to clench the hand when the blow on the tendon was alone given. The hand-response appeared consistently. Furthermore, it showed transfer to the other side, even though the leg on this side had never previously received

[6] As has been demonstrated by Schlosberg, *op. cit.*, 484 f., and by G. R. Wendt, An analytical study of the conditioned knee-jerk, *Arch. Psychol.*, 1930, no. 123, 72–82.

[7] G. R. Wendt, *op. cit.*, 75.

[8] J. M. Stephens, *op. cit.*, 225 f.

a blow in conjunction with the hand-response. The clench-to-the-blow in this thoroughly trained subject appeared even when no ready signal was given. The subject was aware of nothing except that the usual procedure seemed to have gone wrong during these tests.

Hunter and Hudgins [9] have suggested that it might be well to discard the distinction between voluntary and involuntary, and that "voluntary behavior is essentially a conditioned response . . . under the control of self-excited receptor processes." Certainly if the category of the conditioned response is to be a useful one in psychology it must be brought into relationship with that of voluntary action. In disagreement with Hunter and Hudgins, however, we believe that this task may be better served by holding to both concepts until they can be synthesized on grounds yet to be discovered than by discarding one in favor of the other. Whether a voluntary act is a conditioned response or a conditioned response is a voluntary act is less important than the evidence that they are in many ways closely related.

Summary.—If the component events of a voluntary hand-reaction are employed as the conditioned 'stimulus,' the knee-jerk may in some subjects be conditioned rapidly and with stability. The conditioning which has occurred is equally effective in either leg, suggesting that what has been learned is not an isolated response but something similar to a habit or skill.

The result of modifying the complex serving as the conditioned stimulus is to decrease the number of conditioned kicks. The stimulus-complex may be changed considerably without completely destroying its effectiveness.

In the opinion of the writers the difficult problem of how conditioned responses of this type are to be related to voluntary responses, or *vice versa*, remains to be solved.

(Manuscript received August 13, 1934)

[9] C. V. Hudgins, *op. cit.*, 45 ff. See also W. S. Hunter and C. V. Hudgins, Voluntary activity from the standpoint of behaviorism, *Jour. Gen. Psychol.*, 1934, **10**, 198–204.

MACROPHONIC SPEECH

BY E. W. SCRIPTURE

London

In making his parochial rounds the Abbé Rousselot observed that the speech of the people changed gradually from parish to parish. He also noticed that there were successive differences of speech in the three generations of the family of which he was a member, although they all lived in the same house. Even with the most careful listening he could not detect wherein the differences lay. On coming to Paris he sought advice from Gaston Paris. "Only mechanical registration will help you," said the linguist. "Experiments are being made in Marey's laboratory. Go and inquire." As a result of his visit Rousselot devised a method of registering speech that is shown in its latest form in Fig. 1. A funnel-like mouthpiece with a rubber air cushion is placed closely but lightly over the mouth. The currents and puffs of air produce changes of pressure that pass down a wide rubber tube to a membrane of oiled silk whose movements are enlarged and registered on a moving surface of smoked paper. A glass tip inserted into one of the nostrils conducts the changes in pressure to a second recorder. The membrane in the mouth recorder consists of a piece of oiled silk massaged to remove stiffness. It follows the variations of air pressure like a soft air-tight piece of cloth; it is not elastic or vibratory. It is connected to the straw lever by a small piece of soft rubber. The lever is held in a minute axle made of electron; it carries a writing point of parchment. The construction of the lever gives a minimum weight to be moved by the membrane. The nasal recorder has a specially delicate membrane of thin rubber sealed to the edge of the cup by shellac. The lever arrangement is the same as that of the mouth recorder. Special screws provide for the horizontality of the lever and the movement of the recorder longitudinally or sideways.

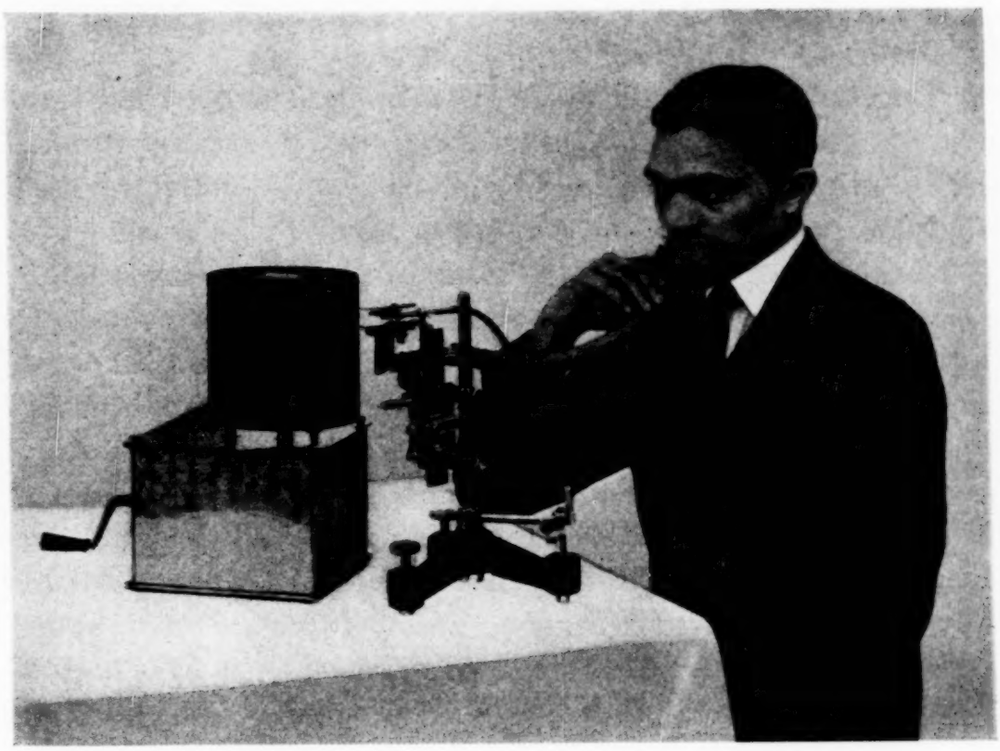

FIG. 1. Making macrophonic registrations of speech.

A registration of the word *fainter* by a double macrophonic recorder is reproduced in Fig. 2. In such registrations the rise of the line indicates an increase in the current of air issuing from the mouth or nose and a fall indicates a decrease. The rate of increase or decrease varies from moment to moment. In the mouth registration in Fig. 2 the increase is

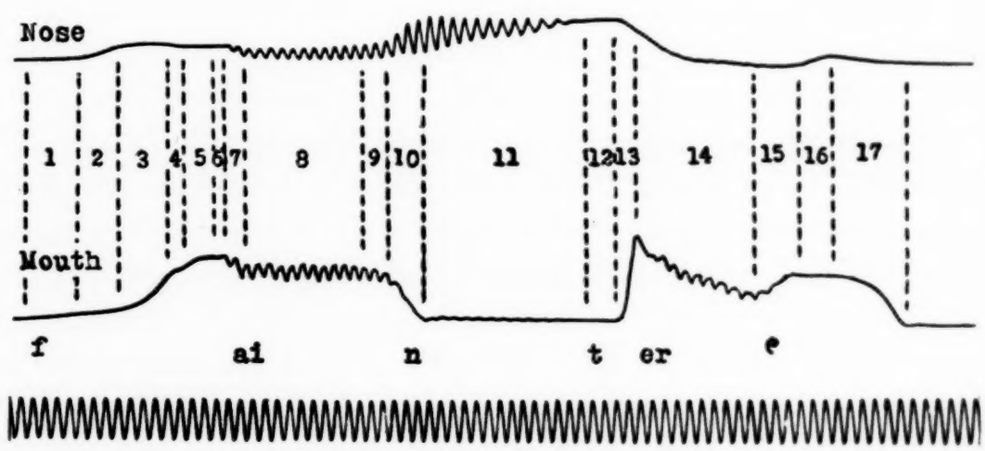

FIG. 2. Macrophonic registration of *fainter*.

at first very slow. During the portion marked 2 the increase is accelerated slightly; during 3 it is still more accelerated. During 4 there is rapid diminution of the rise for an instant. During 5 the rise is again rapid. During 6 there is momentarily neither rise nor fall; there is also an indication of a slight waviness. During 7 there is a fall with strong waves. During 8 the line of waves remains at the same height. During 9 the height remains the same but the waves decrease in amplitude. During 10 the line falls rapidly and the waves become rapidly smaller. During 11 the waves on the base line become gradually fainter. During 12 there is nothing but a stretch of base line. During 13 the breath line shoots up explosively. During 14 the breath line falls steadily with waves. During 15 the line rises again; the waves continue but become weaker. During 16 and 17 there are only changes in the breath line. In the nasal line there is at first no indication of breath. During 2 there is a small rising breath. During 3 the breath remains constant. A slight fall during 4 is followed by a constant current during 5. Faint waviness appears during 6. During 7 the current falls slightly and strong waves appear. During 8 the waves increase slowly in height. During 9 the increase is more rapid and during 10 it rises very rapidly to a maximum while the breath current also increases. During 11 the breath current steadily increases while the waves become gradually weaker. During 12 the breath current remains just as strong or a trifle stronger; no waves are present. During 13 the breath falls rapidly. During 14 the breath current falls gradually to zero. It rises slightly again during 16 and fades away during 17.

Registrations of this kind show that the currents and puffs of breath from the mouth and nose constitute a complete system of speech. As this system is composed of mass movements of air we will term it *macrophonic speech.*

Each variation in the lines in Fig. 2 down to the minutest details registers some fact of macrophonic speech. One of the first facts observed in making such registrations is that no two are ever exactly alike. Each registration of *f* and of every other sound by the same speaker will be slightly dif-

ferent from every other one just as the written signatures of the person's name differ. A macrophonic registration of the person's name by the person himself could be described as a macrophonic signature. The difference between a hand signature and a voice signature lies in the important fact that the voice signature is a very delicate and complicated affair that expresses every shade of peculiarity of the speaker or his condition, whereas the hand with a pen is a clumsy affair. The speaker's personality, his emotional attitude and other factors of his mental and bodily condition express themselves in the registrations. The method is so delicate that several conditions of disease can be detected and diagnosed long before the facts can be ascertained otherwise. Some registrations made with the hope that the method could be applied to determining character for the purposes of psychotechnics showed that the macrophonic registrations by persons of different characters were clearly different. The registrations vary systematically from person to person even in the same family and from one region to the next even with the same dialect. Rousselot had before his eyes in measurable form the two systems of differences. He and later investigators extended the work to the differences among dialects and languages.

Unless the breath current is caught by the apparatus at the mouth and nose it is dissipated immediately in front of the face. It does not pass away from the speaker and it cannot be heard. Rousselot attempted to register what he heard and succeeded in registering something that could not be heard. He discovered, indeed, a system of speech that had not been recognised before. He did not recognise this fact and both he and later investigators have worked in a confusion of ideas concerning the speech that is registered by the graphic method and the speech that can be heard. To have been quite logical Rousselot should have collected registrations on a phonograph. This would not have done him any good because the only thing he could have done with phonograph records would have been to listen to them. This would have been of no advantage because he already had the

opportunity of listening to living speakers as often as he wished and no amount of listening could furnish the information desired.

The speech that is propagated away from the speaker may be heard or may be registered by an apparatus at some distance from the speaker. Its nature can be determined by the registrations on a sound film. It is seen to consist of minute waves that register the vibrations of particles of air. The minuteness of the movements indicates *microphonic speech* as an appropriate term for this form.

The two systems of speech are of different natures. Macrophonic speech consists of mass movements of air; the particles of air pass onward in space. Microphonic speech consists of vibrations of the particles of air around their centers of equilibrium; the vibrations are propagated but the particles do not pass onward in space.

Graphic registration apparatus, ordinary kinematography, X-ray kinematography, stroboscopy, etc. applied to the organs of speech show that they are in constant motion when speech is produced, that they produce the current of breath and that they adjust the sizes, shapes, openings and connections of the vocal cavities. Their movements form a system that can be termed *myokinetic speech.*

In passing through the narrow opening formed between the lower lip and the upper teeth (myokinetic speech) the flat jet of air (macrophonic speech) arouses irregular vibrations (microphonic speech). All three forms are indicated by the symbol *f*. When the opening of the glottis is closed and then opened (myokinetic) a puff of air (macrophonic) is produced that may arouse a vibration in the vocal cavities (microphonic). Such a vibration is termed a vibratory bit. When the puffs are produced in rapid succession, the macrophonic registration shows a series of smooth waves each representing a single puff. The microphonic registration shows a series of adjacent vibratory bits, each representing the result of the action of a puff on the vocal cavities. In macrophonic registrations only the puffs are recorded; their wave forms are always alike and no differences appear for the

different vowels. In microphonic registrations each vowel has its own specific form of wave. The relations among the three speech systems can be stated as follows: myokinetic speech produces macrophonic speech and regulates the character of microphonic speech; macrophonic speech produces microphonic speech; microphonic speech is produced by macrophonic speech and regulated by myokinetic speech. Microphonic speech is not produced by myokinetic speech; the irregular vibrations and the movements in the vibratory bits cannot be produced directly by muscular movements.

The analysis of a macrophonic registration is begun by dividing the line into the portions during which the curve remains of the same nature. Such a series of portions is indicated by the numbers in Fig. 2. For such a portion the term *speech atom* is appropriate. The registration in Fig. 2 consists of 17 speech atoms.

The printed word *fainter* is a symbol for the event of internal (!) speech that resulted in the macrophonic speech registered in Fig. 2. Studies of printed speech have led to the assumption that a word consists of speech sounds that have independent existence in the language spoken. According to this assumption *fainter* consists of six speech sounds: *f, ai, n, t, er.* A macrophonic registration of *f* spoken above would show a mouth line that rises as in the first five atoms in Fig. 2 and is continued as a fall through about the length of the next three atoms. The registration for *ai* spoken alone would show waves like those in atoms 7, 8 and 9 of Fig. 2 lying only slightly above the base line. The line in this stretch in Fig. 2 cannot be the registration of either *f* or *ai* alone; it is evidently the registration of the sum of the last part of *f* and most of *ai*. The last part of *ai* spoken alone does not fall suddenly as in atom 10; the first part of *n* spoken alone does not start high as in this case. Atom 10 is evidently the registration of the last part of *ai* overlapped by the beginning of *n*. The registration of *t* spoken alone would show a descending line lasting through about a length like that of atom 14. A registration of *er* spoken alone would show waves just above the base line. Atom 14 is thus evidently a registration of the end

of *t* and the most of *er* simultaneously. The word ends with a breath—not indicated in print—whose beginning overlaps the end of *er*. The relations of the speech sounds are shown in Fig. 3.

The overlapping cannot be physical. It is impossible to

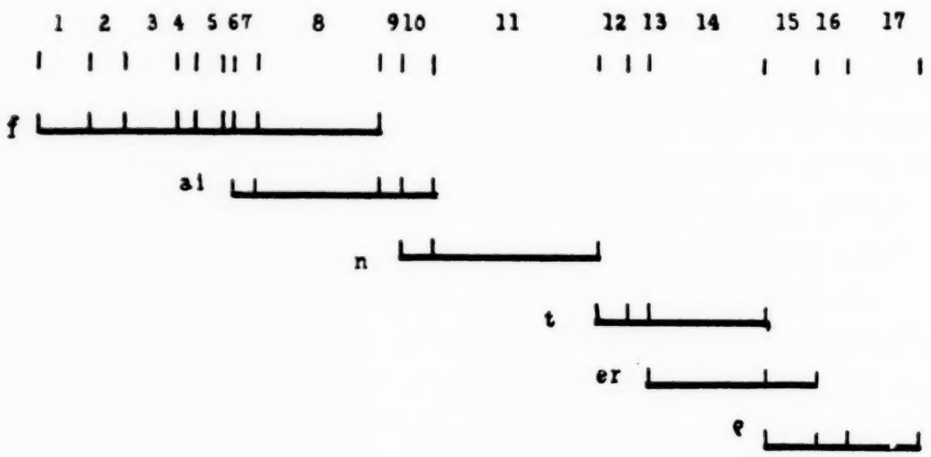

FIG. 3. Overlapping of speech sounds in *fainter*.

speak a series of sounds separately and then push the air movements over one another. The construction of a word by overlapping speech sounds must be a process of internal speech. Internal speech has the dimension of time and various other dimensions corresponding to intensity, quality, etc. The overlapping shows that it has also a dimension in which different speech sounds can exist at the same time. Such a dimension might be termed breadth. Figure 3 would indicate the distribution of the speech sounds in time and breadth.

Having no speech registrations for guidance writers on speech have been misled by printed words to assume that speech consists of a series of sounds just as a printed word consists of a series of letters. They have tried to arrange a 'phonetic alphabet' in which each letter represents a special speech sound. In such an alphabet the word *fainter* would appear as *fente*. Many persons have believed and still believe that the sounds are separated by gaps of silence. Most phoneticians to-day believe that the physical word consists

of a series of speech sounds in which the gaps are filled with transition sounds termed 'glides.' All registrations—macrophonic, microphonic and myokinetic—show, however, that speech consists of continuous movements. There are no constant portions, no blocks of speech and no glides.

THE COMPARATIVE VALUE OF CONTINUOUS VERSUS INTERRUPTED SLEEP

BY RICHARD WELLINGTON HUSBAND

University of Wisconsin

There have been innumerable speculations about the whole problem of sleep. Many people have deplored the fact that one has to sleep, although admitting its necessity. Some feel that it is a pure waste of time. Others think that the hours one sleeps are a matter of habit, and that most of us could cut down materially without harm.

Several have suggested the possibility that one might take his sleep in installments rather in a single stretch, as most of us do now. This latter idea has arisen as a logical assumption from Kohlschütter's work, in which he found that the greatest depth of sleep occurred in the first two hours after falling asleep. Depth was defined as the amount of auditory stimulation necessary to wake one up. A rather intense stimulus was necessary to wake one up after about one-half to two hours after falling asleep, but a small fraction of that amount from then on to the eighth hour sufficed. So it has been suggested that one might sleep two to three hours, stay awake for a number of hours, and then sleep two or three hours more. This procedure would, in theory, take advantage of two stretches of deep slumber.

A possible fallacy in this argument is that the auditory stimulus necessary to awaken one is not necessarily a measure of the benefit being derived from sleep. Because a person can be aroused more easily does not prove that he is deriving less rest and restoration of energy from his sleep.

Miss Helen Rose, a senior major, offered to sacrifice herself physically and socially to carry out a fairly lengthy and rigid routine of sleep regulation to test this hypothesis.

Technique

In planning the experiment, there were several major aspects of technique we considered it desirable to follow. First, since the same individual was to act both as experimental and control subject, it seemed advisable to continue each sleep schedule over fair periods of time. This allowed the subject to fall into more of a regular rhythm than might have been possible if the routine were changed more frequently. She slept one month eight hours a night uninterrupted. The next month had three hours of sleep, three of waking, which were spent in writing, studying, sewing, etc., and three more hours of sleep. The hours of sleep during the experimental month were 11–2 and 5–8. The time was kept very regular, but any minor fluctuations merely shifted the schedule forward or backward, rather than breaking it up.

A thorough physical examination was given at the beginning of the investigation, to make sure that any possible cumulative fatigue would do no harm, and also to compare with later examinations during the experimental period. Several complete physical tests were run during the experiment, and every week certain tests were made. Health was excellent and remained so during the program. One minor cold occurred; this was during the month of normal sleep.

The subject was measured on a large battery of tests each Saturday morning. These involved mental and physical functions, both of speed and accuracy. We especially included two which Laslett found to be most affected after seventy-two hours without sleep: bodily sway and intelligence.

The tests were as follows: 1. Steadiness: measure of muscular tremor by tracing with an electric stylus through two narrow grooves. 2. Speed of tapping. 3. Pursuit rotor: accuracy of eye-hand coördination in following a rotating target with a loosely swivelled pointer. 4. Spool packing: number packed in two minutes. 5. Serial discrimeter: one hundred serial reactions to numbers corresponding to the fingers, each reaction bringing into view the next stimulus. 6. Card sorting: sorting speed for forty cards according to the

'one-behind' method. 7. Code substitution. 8. Psychological examination: various forms of the American Council, Henmon-Nelson, and Alpha. 9. Ataxiameter: cumulative measure of body sway in two minute periods. 10. Dynamometer: strength of grip. Many of these tests are more fully described in articles by Seashore, Miles, and Husband. The physiological tests were: temperature, pulse rate, respiration; blood pressure, metabolism, weight, blood analysis.

The subject, having done considerable laboratory work, was already familiar with practically all the tests, so the gains from practice were not large. She was not, however, at a physiological limit, as is apparent from the data.

Results

In general, it may be stated that there are no consistent differences between interrupted and continuous sleep as shown in our results, either intellectual, motor, or physiological. Since there was only one subject, the differences that do appear may be probably individual fluctuations, rather than systematic trends caused by the experimental conditions. These fluctuations are small, however, no doubt being smoothed out at least partially because of the previous practice of the subject on the tests.

TABLE 1

Results, Coördination Tests

Week	Steadiness	Spool Packing	Pursuit Rotor	Serial Discrimeter*	Card Sorting*
	Control Month				
1	17.9	150	18.0	52.1	54.8
2	18.0	166	50.3	49.6	29.5
3	16.5	156	30.7	53.8	30.5
4	12.0	174	58.3	47.9	26.6
5	14.0	165	91.0	46.2	27.2
	Experimental Month				
6	16.2	170	71.0	46.7	29.1
7	12.5	168	83.0	41.5	25.5
8	14.8	152	75.7	45.5	26.0
9	15.9	160	109.0	42.5	24.7
10	16.2	162	99.3	42.0	23.1

TABLE 2

RESULTS, SIMPLE MOTOR AND MENTAL TESTS

Week	Tapping	Dynamometer	Ataxiameter*	Pulse*	Code Substitution	Intelligence
	Control Month					
1	40.0	35.5	272	76	13.0	96%
2	42.7	37.5	220	83	14.0	98
3	34.7	38.5	246	87	14.0	97
4	39.3	38.5	315	86	15.0	100
5	37.7	38.2	287	90	14.5	97
	Experimental Month					
6	37.7	38.8	441	88	15.5	95
7	32.3	38.0	161	96	16.0	98
8	37.0	37.5	274	72	15.5	100
9	37.7	38.5	259	84	22.0	99
10	35.7	36.5	465	72	23.0	99

* Those tests marked thus have as a goal a low score, being measured in time of error units. The other tests are better as the scores increase.

There are only three tests, of the eleven used, which disclose any signs of deterioration. Steadiness, measured in terms of centimeters traversed before touching the sides of the path, averages lower in the experimental month, although the very lowest score did come during the control month. The two highest tapping speeds came during the control month, but the other three performances during the same month were just about the same as those occurring during the experimental period. In the ataxiameter, sway showed a great deal more variability during the experimental month, and the two highest scores also appeared then.

In contrast to these trends, we have many more where there is no evidence of deterioration at all. In fact, in all the tests where there is an opportunity for learning, there is a great deal of improvement. One may notice especially the pursuit rotor, card sorting, and code substitution. The experimental month does not seem to interrupt the progress of learning at all; scores come down as fast as during the control period, subject of course to the usual limitations of daily fluctuations and diminishing returns.

Other tests, not involving learning, show constant results:

intelligence, dynamometer, and pulse. All physiological measures, blood counts, metabolism tests, etc., remained constant throughout both the control and experimental periods.

Summary

1. To test the influence of interrupted sleep, a subject slept eight consecutive hours nightly for one month, and for the second month three hours, remained awake three, and then slept three more. Eleven mental and motor tests were administered each Saturday morning during both months. A physical examination was also taken each week at the same time.

2. The differences in performance between the two periods were very negligible. In those tests where there was learning, improvement kept taking place in spite of the alteration of sleep routine. Where performance ordinarily remains constant, there appeared only the usual daily fluctuations. There were suggestions as to very slight loss of efficiency in body sway and speed of tapping, but these were slight and inconsistent.

3. Health, from physical examinations and chemical tests, and from the subject's own feelings, remained excellent.

4. It is recognized that the use of only one subject makes the results only suggestive. However, all indications are in one direction, and we may say with complete assurance that the subject suffered no ill consequences at all from the slight reduction of sleep and the complete change in normal sleep habits.

(Manuscript received August 8, 1934)

References

1. Husband, R. W., and Ludden, M. J., Sex differences in motor skills. J. Exp. Psych., 1931, 14, 414–422.
2. Husband, R. W., and Miles, W. R., Card sorting by reaction to the previous card. *J. App. Psych.*, 1927, 11, 465–482.
3. Laslett, H. R., Experiments on the effects of the loss of sleep. J. Exp. Psych., 1928, 11, 370–396.
4. Seashore, R. H., Individual differences in motor skills. *J. Gen. Psych.*, 1930, 3, 38–66.
5. Seashore, R. H., Stanford motor skills unit. *Psych. Monog.*, 1928, 39, 51–66.

DISTRIBUTION OF EFFORT IN A PERCEPTUAL–MOTOR TASK

BY E. B. SKAGGS

Wayne University

In this investigation we have attempted to ascertain the relative efficiency of three methods of distributing practice in a perceptual-motor task. The task consisted in replacing the blocks, which were in a standard position at the beginning of every trial, in a Goddard formboard. We regard the task as one involving a simple type of perceptual-motor learning, a type of learning which results in a rather typical learning curve showing negative acceleration and one or more plateaus. Efficiency was measured in terms of the average amount of time needed to replace the blocks. Hence, all figures which appear in the following tables express time in seconds and fractions thereof. Efficiency was measured at the end of the period of learning and again four weeks after the completion of the learning.

Apparatus and Procedure

The apparatus involved was very simple. Two Goddard formboards were used, one to hold the blocks at the beginning of every trial, the other to receive the blocks when the subject replaced them. The first board we shall refer to as the container or C-board. The second we shall call hereafter the receiving or R-board. The table upon which the boards were placed was 30½ inches high. The R-board was screwed down to the table along the edge. The C-board was likewise screwed down on the table just directly beyond the R-board. The top edge of the R-board was in contact with the bottom edge of the C-board. The C-board was so placed that the recesses corresponded spatially to the recesses of the R-board. The recesses of the C-board were partly filled with cardboard so that every block projected ¾ inch above the level of the board. This gave ample space for grasping the blocks in every case. At the beginning of every trial the blocks were placed in their proper recesses in the C-board.

The experiment was done in the laboratory dark-room under conditions of constant artificial light. A 200-Watt light, with an extended screen on the side of the subject so that the light-rays would not shine into her eyes, was suspended 40 inches above and somewhat to the front of the table. The temperature averaged about 72 degrees Fahrenheit with a variation of some two degrees. The room was very quiet and only the subject and experimenter were present during the experiment. The

writer, who was the experimenter in all cases, used a split-seconds stop-watch to record the time of every trial.

In every case the following procedure was followed. At the beginning of every trial the experimenter stood in front of the subject and directly in front of the R-board. Thus, the experimenter's body was a screen to hide the 'set-up' from the subject. At the beginning of every trial the experimenter said, "All right. Get ready." Then he stepped to the right, saying aloud in slow rhythmical tone, "One, Two, Go." As he said 'Go' the experimenter started his stop-watch.

As the experimenter stepped aside the subject stepped up into position, bent over the boards with hands almost touching the two blocks which were to be replaced first, and at the signal 'Go' began to replace the blocks from the C- to the R-board. The subject continued, using both hands and replacing two blocks at a time in a prescribed order, until all blocks were replaced. The prescribed order for replacing the blocks was as follows: *first*, diamond, star; *second*, circle, oval; *third*, oblong, square; *fourth*, cross, hexagon; and *fifth*, triangle, semi-circle. These blocks were numbered 1,1, 2,2, 3,3, etc., the numbers being printed in bold face upon every pair of blocks.

As the last block was replaced the experimenter stopped his watch, noted the time in seconds and fractions thereof, informed the subject of the time taken for the trial, and recorded said time. The blocks were then replaced in the farther C-board, ready for the next trial. When the subject replaced the last block she stepped back to her original position as the experimenter stepped back in front of the boards. After completing a trial, the subject at once noted the position of the seconds-marker on an Eastman timer which was placed on a shelf to the front and right. At the end of 25 seconds the subject informed the experimenter that the rest-time was ended. We asked our subjects to keep the time for the rest-periods in order to keep them from thinking of the task at hand and also to keep their attention from drifting to extraneous matters. This tended to standardize the conditions of the rest-period. Needless to say the experimenter checked upon every subject occasionally and the subject knew that she was being checked from time to time. The actual time which elapsed between every trial was approximately 30 seconds.

The first meeting with every subject occurred on a Friday. At this time the nature of the work was explained. The subject was instructed to work to improve speed, to eliminate waste movements, and to become just as expert as possible in the time allotted for training. She was asked to avoid thinking of the task when away from it. The subject was told that she would be recalled four weeks later, after the completion of the regular training period, for a four-weeks retention test.

At this first meeting the experimenter worked the board once for the subject while she observed. Then the subject worked the board two times slowly and carefully without being timed. There then followed ten trials every one of which was timed by the experimenter.

Regular work was begun on the following Monday. Every subject was assigned to one of three groups. Group A worked the board *twenty times* every day for five consecutive days. Group B worked the board *ten times* every day for ten consecutive days, omitting Saturday and Sunday. Group C worked the board *five times* every day for twenty consecutive days, omitting Saturdays and Sundays. This made, in all, a total of 112 times that every subject worked the formboard. At the completion of the regular training period every subject was asked to report exactly four weeks later for a retention and relearning test. Exactly four weeks later every subject reported and worked the board *ten times*.

Our subjects were all college women of sophomore standing. Twenty subjects constituted a group. The subjects were selected at random and, as Table I below

TABLE I

PRESENTING A COMPARISON OF INITIAL FORMBOARD ABILITY OF GROUPS A, B, AND C. BASED UPON THREE CRITERIA

Criterion Used		Group A	Group B	Group C
First 10 practice trials...	Mean	6.04	5.90	6.09
	M.V.	0.77	0.45	0.73
	S.D.	0.89	0.54	0.83
First 10 regular trials....	Mean	5.22	5.07	5.38
	M.V.	0.57	0.20	0.50
	S.D.	0.73	0.27	0.69
First 20 regular trials....	Mean	4.94	4.90	5.17
	M.V.	0.51	0.17	0.44
	S.D.	0.63	0.21	0.63

shows, the three groups were of approximately equal initial formboard ability. This is shown to be true whether one judges efficiency by the first ten practice trials, the first ten regular trials, or the first twenty regular trials. According to the results of the Otis Self-Administering Tests the groups were of approximately the same intellectual level.

RESULTS

The writer carried out this experiment as a result of curiosity to find out whether a greater distribution of practice would be more effective than a more massed form of distribution in the case of perceptual-motor learning of the kind already described. The results of the investigation have been somewhat surprising to him.

Table II shows the results of applying the three different methods of training to the three groups of subjects. These three groups were comparable in mental ability, initial formboard ability, general health, sex, and chronological age.

The data of Table II show that Group A, using the more massed form of practice, improved most by all three criteria used. Group C, the group that used the most distributed form of practice, improved least. The differences appear to be persistent and reliable.

In Table III we present the results of the four-week retention (or retention-relearning) tests of the three groups.

TABLE II

PRESENTING A COMPARISON OF GROUP AVERAGES AND MEASURES OF VARIABILITY FOR GROUPS A, B, AND C

Criterion Used		Group A	Group B	Group C
Last 20 trials.....	Mean	3.85	4.01	4.12
	M.V.	0.25	0.25	0.34
	S.D.	0.29	0.32	0.42
Last 10 trials.....	Mean	3.78	3.98	4.02
	M.V.	0.32	0.22	0.35
	S.D.	0.40	0.25	0.46
Last 5 trials.....	Mean	3.76	3.98	4.02
	M.V.	0.28	0.33	0.32
	S.D.	0.36	0.43	0.41

TABLE III

SHOWING GROUP AVERAGES AND MEASURES OF VARIABILITY ON THE FOUR-WEEK RETENTION TEST

	Group A	Group B	Group C
Mean.....	3.86	3.92	4.03
M.V.....	0.36	0.19	0.31
S.D.....	0.45	0.49	0.38

The averages are based upon the averages for the ten trials of every subject.

The figures of Table III indicate clearly that the more massed form of practice results in the greatest improvement. Group A is shown to be the most efficient, Group B is next most efficient, and Group C, with most distributed practice, is least efficient.

CONCLUSION

Under the conditions of this experiment, in which college women served as subjects, a method of more massed practice has proved to be more effective than the more distributed methods of practice. The more massed form of practice also resulted in greater efficiency in the four-week retention or relearning test.

The above findings merely emphasize a fact which every experimental psychologist must accept and face, namely, the

fact that it is impossible to generalize in psychology. A given result or fact holds only for a given set of specific conditioning factors. Ruch,[1] after a careful inventory of past work on distribution of effort, has adequately pointed out the classes of conditioning factors which must be considered.

(Manuscript received August 3, 1934)

[1] Ruch, T. C. *Psychol. Rev.*, 1928, **35**, 19–45.

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Herbert S. Langfeld, Editor PSYCHOLOGICAL REVIEW,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Samuel W. Fernberger, Editor JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
University of Pennsylvania, Philadelphia, Pa.

Contributions intended for the Psychological Monographs should be addressed to

Professor John F. Dashiell, Editor PSYCHOLOGICAL MONOGRAPHS,
University of North Carolina, Chapel Hill, N. C.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor John A. McGeoch, Editor PSYCHOLOGICAL BULLETIN,
Wesleyan University, Middletown, Conn.

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor PSYCHOLOGICAL INDEX,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company, Princeton, New Jersey

The

Journal of Experimental Psychology

is indexed in the

International Index to Periodicals

to be found in most public and college libraries

AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, Madison Bentley, K. M. Dallenbach, and E. G. Boring.
Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per yr.; $7.00 per vol. 1,000 pages ann. (2 vols.). Edited by Carl Murchison.
Quarterly. Child behavior, animal behavior, comparative psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 540 pages annually. Edited by Herbert S. Langfeld.
Bi-monthly. General psychology. Founded 1894.

Psychological Monographs—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per vol. 500 pages. Edited by John F. Dashiell.
Without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.; Psychological Review Company.
Subscription $4.00. 400–500 pages. Edited by Walter S. Hunter and R. R. Willoughby.
An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company.
Subscription $6.00. 720 pages annually. Edited by John A. McGeoch.
Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N. Y.; Columbia University.
Subscription $6.00. 500 pages per volume. Edited by R. S. Woodworth.
Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $5.00. 448 pages annually. Edited by Henry T. Moore.
Quarterly. Abnormal and social. Founded 1906.

Psychological Clinic—Philadelphia, Pa.; Psychological Clinic Press.
Subscription $3.00. 288 pages. Edited by Lightner Witmer.
Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.

Journal of Educational Psychology—Baltimore: Warwick & York.
Subscription $6.00. 720 pages. Monthly except June to August.
Edited by J. W. Dunlap, P. M. Symonds and H. E. Jones. Founded 1910.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W.
Subscription $6.00. 500 pages annually. Edited by W. A. White and S. E. Jelliffe.
Quarterly. Psychoanalysis. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J.; Psychological Review Company.
Subscription $7.00. 900 pages annually. Edited by Samuel W. Fernberger.
Bi-monthly. Experimental psychology. Founded 1916.

Journal of Applied Psychology—Indianapolis; C. E. Pauley & Co.
Subscription $6.00. 600 pages annually. Edited by James P. Porter, Ohio University, Athens, Ohio. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Company.
Subscription $5.00 per volume of 450 pages. Ed. by Knight Dunlap and Robert M. Yerkes. Two volumes a year. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; The Johns Hopkins Press.
Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor.
Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press.
Subscription $7.00 per vol. One volume per year. Edited by Carl Murchison.
Bi-monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $6.00. 700 pages ann. Edited by Walter S. Hunter and R. R. Willoughby.
Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per yr.; $7.00 per vol. 1,000 pages ann. (2 vols.). Edited by Carl Murchison.
Quarterly. Experimental, theoretical, clinical, historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.; Clark University Press.
Subscription $7.00. 500 pages annually. Ed. by John Dewey and Carl Murchison.
Quarterly. Political, racial, and differential psychology. Founded 1929.